

VISIT...

LANZAROTE
Caliente.COM

*Р. Рихтмайер,
К. Мортон*

РАЗНОСТНЫЕ МЕТОДЫ
РЕШЕНИЯ
КРАЕВЫХ ЗАДАЧ

*Перевод со второго английского издания
Б. М. БУДАКА, А. Д. ГОРБУНОВА,
В. Е. КОНДРАШОВА и В. Е. ТРОЩЕВА*

*Под редакцией
Б. М. БУДАКА и А. Д. ГОРБУНОВА*

ИЗДАТЕЛЬСТВО «МИР»
Москва 1972

Второе, существенно расширенное и переработанное издание одноименной книги первого из авторов. Первое издание также было переведено на русский язык (ИЛ, 1960).

Книга посвящена разностным методам решения задачи Коши и смешанной задачи для уравнений в частных производных. В ней рассматриваются не только вопросы теории, но и большое количество конкретных задач, имеющих важное практическое значение (уравнение теплопроводности, волновое уравнение, уравнения газовой динамики, уравнение переноса и др.).

Книга интересна для математиков, занимающихся теоретическими вопросами вычислительной математики, для специалистов по дифференциальным уравнениям, для механиков, физиков и инженеров, занимающихся приложениями разностных методов к решению конкретных задач. Доступна студентам старших курсов и аспирантам указанных специальностей.

Редакция литературы по математическим наукам

ОТ РЕДАКТОРОВ ПЕРЕВОДА

Разностные методы решения краевых задач уже давно привлекают внимание математиков как в Советском Союзе, так и за рубежом; в разработке и применении этих методов немалую роль сыграли труды советских ученых.

В последнее время в связи с развитием новых областей техники и появлением быстродействующих электронных цифровых машин эти методы развиваются особенно быстро, публикуется большое количество работ, целиком или частично посвященных разностным методам. В этой связи следует упомянуть книги О. А. Ладыженской «Смешанная задача для гиперболического уравнения», В. С. Рябенского и А. Ф. Филиппова «Об устойчивости разностных уравнений», С. К. Годунова и В. С. Рябенского «Введение в теорию разностных схем», А. А. Самарского «Введение в теорию разностных схем», Н. Н. Яненко «Метод дробных шагов для решения многомерных задач математической физики», И. С. Березина и Н. П. Жидкова «Методы вычислений», а также ряд важных статей, посвященных разностным методам решения краевых задач.

В обширной литературе о разностных методах книга Рихтмайера и Мортонa занимает несколько своеобразное место.

В ее первой, теоретической части излагается общая теория разностных методов решения уравнений в частных производных — устойчивости и сходимости соответствующих разностных схем. Во второй части полученные результаты применяются к конкретным физическим задачам. Это — задачи переноса (нейтронов и фотонов), одномерные и многомерные задачи газодинамики, задачи теплопроводности и диффузии, а также задачи о колебаниях струн и стержней.

За годы, прошедшие после выхода первого издания, были созданы новые вычислительные машины и разработаны новые методы численного решения задач. В связи с этим при подготовке нового издания книги текст ее был существенно дополнен и несколько переработан. Так, в первой ее части читатель найдет новые главы, посвященные уравнениям с переменными коэффициентами и общей теории смешанных краевых задач; во второй части значительно расширена глава о числен-

ном решении одномерных задач газовой динамики и добавлена глава о расчете многомерных течений.

Новые главы и параграфы книги переведены В. Е. Кондрашовым и В. Е. Трощевым. В основу русского текста остальных глав, мало измененных в оригинале, положен перевод первого издания, выполненный авторами этих строк.

Расширенная по сравнению с первым изданием библиография все же недостаточно отражала работы советских ученых; чтобы восполнить этот пробел, для настоящего издания специально составлена дополнительная библиография.

Мы надеемся, что эта книга, как и первое ее издание, заинтересует широкий круг читателей — от студентов до научных работников и окажется полезным дополнением к отечественной литературе по численным методам решения краевых задач.

Б. М. Будаков, А. Д. Горбунов

ПРЕДИСЛОВИЕ КО ВТОРОМУ ИЗДАНИЮ

За десять лет, прошедшие со времени выхода в свет первого издания этой книги, число работ, посвященных теории различных методов для уравнений с частными производными, возросло на несколько порядков, и читатель мог бы предположить, что основная цель переиздания книги состоит в изложении новых разделов теории (или той их части, которая относится к решению краевых задач), однако это не было для нас главной целью. Хотя в некоторых главах, особенно в гл. 4—6, весьма подробно излагается ряд важных теоретических результатов, книга предназначена прежде всего для тех, кто занимается решением конкретных прикладных задач, а не для специалистов по численному анализу. Как и в других частях математики, абстрактная теория и приложения имеют здесь тенденцию развиваться независимо друг от друга, и в результате этого развития в чистой математике создано новое направление. Хотя это новое направление уже привело к многочисленным результатам и понятиям, имеющим значение для прикладников, практические задачи физики в большинстве своем остаются пока в том или ином отношении слишком сложными и не покрываются доказанными теоремами. При проведении практических расчетов все так же необходимо сочетать физическую интуицию с соображениями эвристического характера и численными экспериментами, как это было и четверть века назад, когда Дж. фон Нейман положил начало развитию этого направления.

Мы включили в книгу некоторые разделы теории, которые по нашему мнению находят применение или имеют значение для практических приложений, но постарались также сохранить тот практический подход, который был характерен для первого издания. Основные теоретические достижения таковы: 1) завершено создание теории для однородных краевых задач с постоянными коэффициентами, а именно Баченаном и Крайсом для таких задач общего вида получены достаточные условия устойчивости, и 2) построена строгая теория устойчивости для определенных классов задач с переменными коэффициентами, смешанных краевых задач и квазилинейных задач. Для лиц, занимающихся решением практических задач, на наш взгляд представляет ин-

интерес следующее: 1) понятие диссипативных разностных схем; 2) метод Лакса — Вендроффа для систем уравнений с законами сохранения; 3) методы чередующихся направлений для многомерных параболических задач; 4) практические критерии устойчивости для тех случаев, когда полученных обычным путем условий устойчивости оказывается недостаточно, и 5) применение энергетических методов для анализа устойчивости.

В современных теоретических исследованиях обычно используются такие математические понятия и методы, которые мало знакомы либо вообще недоступны для лиц, не являющихся специалистами по численному анализу; поэтому для нас было трудной проблемой изложение нового материала — многим читателям доказательства все же покажутся сложными, несмотря на все наши попытки упростить их. Насколько мы понимаем, это неизбежно, но по нашему мнению включение доказательств оправдывается важностью результатов.

Мы хотим выразить особую признательность Хайнцу Крайсу, стимулировавшему нашу работу, и Питеру Лаксу, Гильберту Стрэнгу, Бересфорду Парлетту и Сэму Шехтеру за полезные обсуждения. Мы благодарны также Валери Бриз и Дженни Эйплэнд за большую помощь в сложном деле оформления математической рукописи.

ПРЕДИСЛОВИЕ К ПЕРВОМУ ИЗДАНИЮ

Совершенно очевидно, что появление современных вычислительных машин должно привести к своего рода революции в области численных методов. Сейчас численные методы применяются там, где лет пятнадцать назад об их применении и не помышляли. Хотя электронные вычислительные машины выполняют в основном те же операции, которые производились при ручном счете, однако скорость и емкость этих машин позволяют им осуществлять процедуры, совершенно недоступные для ручного счета.

Конечноразностные методы решения уравнений в частных производных были предложены еще в 1928 г. в знаменитой статье Куранта, Фридрихса и Леви, но на практике эти методы стали применяться лишь 15 лет спустя, и стимулом к их практическому использованию послужило быстрое развитие техники во время войны, а также возможность применения первых автоматических вычислительных машин. В Лос-Аламосе, например, в работе лаборатории в военное время существенное место занимали расчеты некоторых нестационарных течений; этим расчетам, производившимся на существовавших тогда машинах с перфокартами, уделялось много времени и усилий. Сразу после войны члены Лос-Аламосского персонала в Абердине численно решили на машине ЭНИАК еще более сложные задачи, относящиеся к системам уравнений в частных производных, а затем в разных местах США были решены многие задачи, связанные с движением жидкостей и газов, диффузией и переносом нейтронов, переносом лучистой энергии, термоядерными реакциями и т. п.

Однако ожидаемая революция в численных методах происходит довольно медленно. Появились некоторые новые идеи, но основные методы по существу остались теми же, что и много лет назад. Гауссов метод исключения (в той или иной форме) все еще остается одним из лучших методов решения систем линейных уравнений, а метод Рунге — Кутта — одним из лучших для решения обыкновенных дифференциальных уравнений. Похоже на то, что эти основные методы классического численного анализа в течение некоторого времени останутся наиболее ходо-

выми и, быть может, они скорее будут включены в методы, которые появятся в будущем, нежели заменены ими. Однако совершенно ясно, что быстродействующим вычислительным машинам с большой памятью по самой их природе присуща способность совершать такие необычайно разнообразные дела, о которых раньше нельзя было и мечтать. Вероятно, в будущем откроются многие мощные методы как для тех задач, которые мы сейчас умеем решать, так и для тех, которые пока представляются неразрешимыми. Методы, которыми мы располагаем в настоящее время, по большей части представляют собой простое объединение методов, придуманных для ручного счета и задач с малым объемом вычислений. Мы, безусловно, даже еще и близко не подошли к использованию всех огромных возможностей вычислительных машин.

В численном анализе наибольшие изменения претерпели, пожалуй, разностные методы решения дифференциальных уравнений в частных производных, особенно уравнений, связанных с нестационарными физическими процессами — они и являются основным предметом данной книги. Эти уравнения для реальных физических задач зачастую оказываются весьма сложными; они имеют переменные коэффициенты и в ряде случаев нелинейны. Часто встречаются смешанные системы (состоящие, скажем, из уравнений гиперболического и параболического типов), задачи с несколькими пространственными переменными и задачи, включающие интегро-дифференциальные уравнения. Число новых алгоритмов для решения таких задач быстро возрастает.

Развитие численных методов, предназначенных для работы на автоматических вычислительных машинах, в одном важном отношении отличается от предшествующего развития численных методов для ручного счета. Новое направление численных методов основывается в первую очередь на эмпирических данных и интуитивных соображениях и в меньшей мере (по сравнению с классическими методами численного анализа) обосновано математически. Это вызвано усложнением решаемых задач и быстрым накоплением алгоритмов для их решения, а также тем обстоятельством, что люди, работающие на электронных вычислительных машинах, вообще говоря, не являются специалистами по классическому численному анализу. Но не следует осуждать за это новое направление в развитии численных методов: ожидание доказательств сходимости и получения оценки погрешности для новых методов надолго задержало бы внедрение большинства вычислительных устройств, используемых сейчас техникой и промышленностью. Ведь по мере того, как усложняются решаемые задачи, трудность доказательств сходимости и определения оценки погрешности быстро возрастает. Поэтому в новом направлении существует известный разрыв

между математиками, занимающимися данными вопросами, и теми, кто практически использует эти методы (например, физиками). В то же время я убежден, что революция, которой мы ждем, осуществится только при совместной работе математиков и практиков. Для уменьшения разрыва между ними нужно, чтобы математики не проявляли излишней ортодоксальности в принятии эмпирического интуитивного подхода к вопросу и чтобы практики придавали большее значение действительному пониманию существа используемых ими методов.

Осенью 1953 г. доктор Лесли Пек и автор настоящей книги, посетившие Институт математических наук Нью-Йоркского университета, по просьбе проф. Куранта организовали семинар, посвященный тем разделам численного анализа, которые представляют интерес как для математиков, так и для вычислителей. В работе семинара участвовали многие сотрудники и гости Института: Дж. Кёртис, Дж. Франклин, У. Гивенс, Е. Изаксон, Ф. Джон, П. Лакс, М. Роуз и другие. Затем в Нью-Йоркском университете были организованы аналогичные семинары и курсы, в которых работали многие новые участники: большая часть профессорско-преподавательского состава Института и персонала входящего в него вычислительного центра, а также некоторые лица, прикомандированные к нему, в частности Дж. Тодд, О. Тодд и Дж. Форсайт.

Настоящая книга содержит часть тех сведений по численному анализу, которые были получены за последнее десятилетие в Лос-Аламосе и ряде других мест и соответственно модифицированы на вышеупомянутых семинарах и курсах. Теоретическая часть построена на идеях Дж. фон Неймана, П. Лакса и отчасти автора этой книги, однако ясно, что при другом подходе к вопросу вклад других исследователей был бы представлен более рельефно. Книга представляет собой попытку объединить математический и практический подходы, и хотя революция в области численных методов по-видимому остается столь же далекой, как и ранее, можно надеяться, что в действительности мы медленно, но верно приближаемся к знанию того, что следует делать с современными вычислительными машинами и как понимать получаемые на них результаты.

Р. Д. Рихтмайер

Нью-Йоркский университет

Общая теория

Глава 1

ВВЕДЕНИЕ

§ 1.1. Краевые задачи

В этой книге рассматриваются краевые задачи, возникающие в различных разделах физики сплошных сред, таких, как теория теплопроводности, диффузия, гидродинамика, магнитная гидродинамика, акустика, электромагнетизм, волновая механика, перенос лучистой энергии, диффузия нейтронов, упругие колебания. Описываемые этими теориями процессы нестационарны и приводят к уравнениям в частных производных или интегро-дифференциальным уравнениям, в которых одна из независимых переменных, обозначаемая через t , играет роль времени. Природа этих уравнений такова, что если произвольно задано состояние физической системы в некоторый начальный момент времени $t = t_0$, то решение существует при $t \geq t_0$ и определяется единственным образом граничными или какими-нибудь другими дополнительными условиями.

Непосредственно темой этой книги являются конечноразностные методы для получения численного решения задач такого рода. Первая часть посвящена общему обсуждению этих методов для линейных задач в рамках теории линейных операторов; цель этого обсуждения — выкристаллизовать основные понятия, связанные с устойчивостью и сходимостью. Для однородных краевых задач с постоянными коэффициентами выводятся результаты общего характера; в последующих главах приводится ряд результатов для некоторых специальных типов уравнений с переменными коэффициентами и для некоторых смешанных краевых задач.

Во второй части описываются основные конечноразностные методы, применяемые для решения краевых задач из различных областей прикладной физики на цифровых вычислительных машинах. Насколько возможно изложение базируется на теоретическом материале первой части и на некоторых опубликованных работах, но поскольку существующая математическая теория пока еще несовершенна (в основном потому, что не охватывает нелинейных уравнений), то мы также призываем на помощь интуицию и эксперимент.

§ 1.2. Уравнение теплопроводности

В этой главе основные идеи иллюстрируются на общеизвестном примере линейной задачи одномерной теплопроводности или диффузии. Если x обозначает координату вдоль тонкого изолированного стержня, по которому может распространяться тепло, или координату, перпендикулярную широкой плите, у которой все точки каждой из ее сторон имеют одинаковую температуру, и если $u = u(x, t)$ обозначает температуру в точке x в момент времени t , то эта температура удовлетворяет дифференциальному уравнению

$$a \frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \left(K \frac{\partial u}{\partial x} \right), \quad (1.1)$$

где $a = a(x)$ или $a(x, t)$ — удельная теплоемкость материала (на единицу объема), а $K = K(x)$ или $K(x, t)$ — коэффициент теплопроводности.

Это уравнение является линейным, хотя оно было бы нелинейным, если бы мы допустили, что величины a и K , характеризующие свойства материала, зависят не только от x и t , но и от температуры u . В этой главе мы ограничимся случаем однородной среды, когда a и K постоянны, и положим $K/a = \sigma > 0$.

Дифференциальное уравнение (1.1) надо дополнить граничными условиями, скажем в точках $x = 0$ и $x = L$, соответствующих концам стержня или сторонам плиты; например, $u = u_0$ при $x = 0$, если этот конец стержня находится в контакте с большим тепловым резервуаром температуры u_0 , или $\partial u / \partial x = 0$ при $x = 0$, если этот конец теплоизолирован.

В общем случае рассматриваемая задача включает в себя одно или несколько уравнений в частных производных или интегро-дифференциальных уравнений вместе с граничными и начальными условиями. Возможны также внутренние граничные условия; например, если $K = K(x)$ имеет разрыв первого рода на границе двух сред в указанной выше задаче, то из физических соображений требуется непрерывность u и $K \partial u / \partial x$ на этой границе. Аналогично на поверхности разрыва в потоке сжимаемой жидкости дифференциальные уравнения гидродинамики заменяются условиями Ренкина — Гюгонио (см. гл. 12).

Мы не будем входить в подробное обсуждение важных, но трудных вопросов существования и единственности решений; задачи, с которыми нам придется иметь дело, будут обычно сформулированы с самого начала в таком виде, чтобы были обеспечены существование и единственность решения при физически приемлемых допущениях (таких, как кусочная гладкость коэффициентов и начальных данных).

При численном решении краевых задач возникают вопросы о построении системы конечноразностных уравнений, о мето-

дах ее решения, о ее устойчивости и ее точности. В нескольких ближайших параграфах настоящей главы эти вопросы будут рассмотрены на примере простейшей задачи теплопроводности, поставленной выше. При этом мы сделаем следующие дополнительные упрощающие предположения: выберем единицу длины так, чтобы $L = \pi$, и будем считать, что граничные условия имеют вид $u = 0$ на концах и что начальная функция $\varphi(x)$ обладает столь высокой гладкостью, какая нам только будет удобна. Наша краевая задача запишется в виде

$$\begin{aligned} u &= u(x, t), \\ \frac{\partial u}{\partial t} &= \sigma \frac{\partial^2 u}{\partial x^2}, \quad \text{где} \quad \sigma = \text{const} > 0, \\ 0 &\leq x \leq \pi, \\ t &\geq 0, \end{aligned} \quad (1.2)$$

$$u(x, 0) = \varphi(x) \text{ (заданная функция) при } 0 \leq x \leq \pi, \quad (1.3)$$

$$u(0, t) = u(\pi, t) = 0 \quad \text{при } t \geq 0. \quad (1.4)$$

Для иллюстрации можно представить себе тонкий плоский горизонтальный электропроводящий слой, заключенный между двумя толстыми плоскими слоями обычного покрытия и нагреваемый проходящим через него электрическим током. При этом предполагается, что нижняя поверхность нижнего покрытия и верхняя поверхность верхнего покрытия имеют комнатную температуру, которую при соответствующем выборе начала отсчета можно считать равной нулю. Если после установления стационарного распределения температур электропроводящий слой внезапно удалить, мы получим для слоев покрытия упомянутую выше задачу с начальной функцией

$$\varphi(x) = \begin{cases} Cx & \text{при } 0 \leq x \leq \frac{\pi}{2}, \\ C(\pi - x) & \text{при } \frac{\pi}{2} \leq x \leq \pi, \end{cases} \quad (1.5)$$

где x — вертикальная координата, а C — некоторая постоянная. Эта функция и типичные распределения температуры $u(x, t)$ в последующие моменты времени изображены на рис. 1.1 сплошными линиями.

Для сравнения с результатами вычислений, производимых при помощи разностных уравнений, заметим, что точное решение рассматриваемой краевой задачи может быть получено известным методом рядов Фурье. Его можно записать в виде ряда по синусам. Если положить $\varphi(x) = -\varphi(-x)$ на отрезке $-\pi \leq x \leq 0$, то точное решение нашей задачи на отрезке

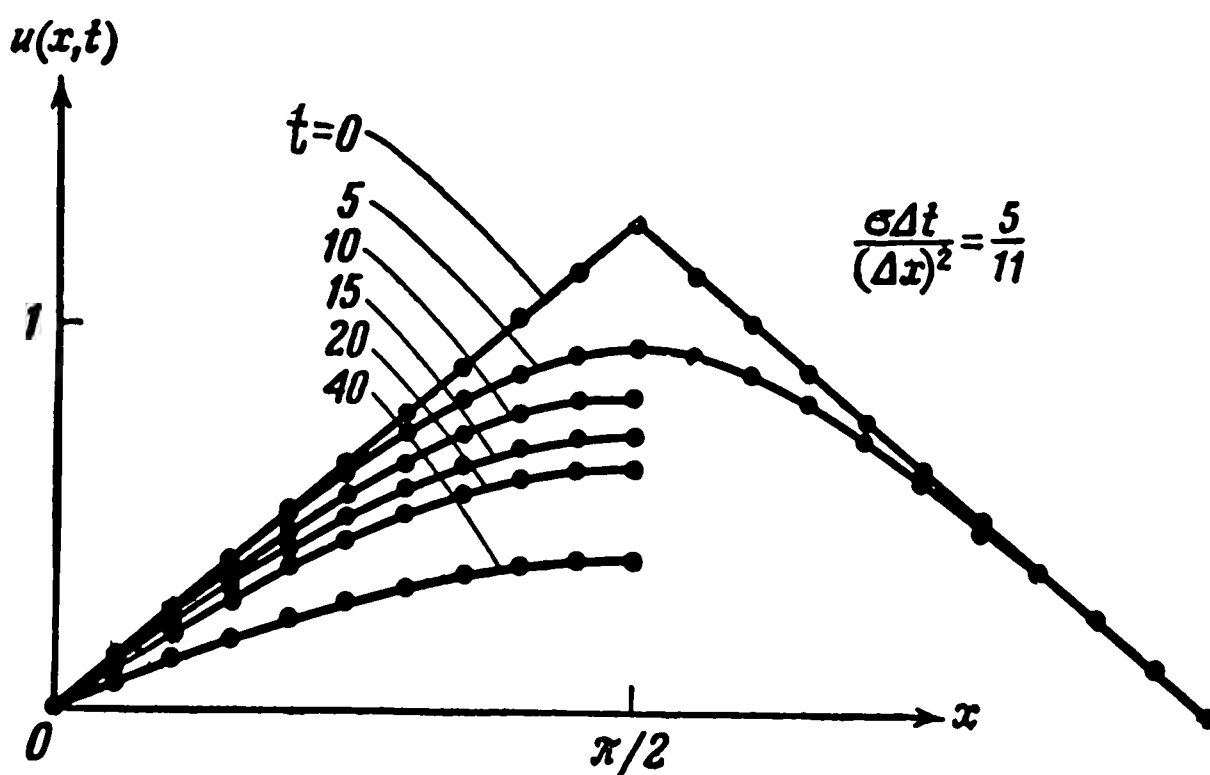
$-\pi \leq x \leq \pi$ можно представить в виде ряда

$$u(x, t) = \sum_{m=-\infty}^{\infty} A_m \exp(imx - m^2 \sigma t), \quad (1.6)$$

где

$$A_m = \frac{1}{2\pi} \int_{-\pi}^{\pi} \varphi(x) \exp(-imx) dx. \quad (1.7)$$

Иногда ряд (1.6) рассматривают как обобщенное решение задачи, если этот ряд сходится лишь при $t > 0$; однако в этой главе мы предположим, что сходится и ряд Фурье для функции $\varphi(x)$, и притом абсолютно.



Р и с. 1.1. Решение одномерной задачи теплопроводности. Кривые соответствуют точному решению, а точки — решению конечноразностных уравнений (1.9)—(1.11). Поскольку все кривые симметричны, на чертеже приведены только половины их.

Для начальной функции (1.5), соответствующей задаче с электропроводящим слоем, мы имеем

$$A_m = \begin{cases} 0 & \text{при четном } m, \\ \frac{2iC}{\pi m^2} (-1)^{(m+1)/2} & \text{при нечетном } m. \end{cases} \quad (1.8)$$

§ 1.3. Конечноразностные уравнения

Очевидно, метод рядов Фурье применим только к весьма ограниченному классу задач. Действительно, успех его применения, по крайней мере в изложенном выше виде, связан с линейностью дифференциальных уравнений, постоянством их коэффициентов и тем фактом, что граничные условия эквивалентны условиям нечетности и периодичности начальных данных.

Это справедливо также и для других уравнений в частных производных, которые мы будем рассматривать, и для задач с бóльшим числом зависимых и независимых переменных. Напротив, применение методов конечных разностей несколько не связано такими ограничениями, хотя анализ их устойчивости и сходимости, который приводится ниже, строго говоря, справедлив лишь при высказанных ограничениях.

Пусть Δx и Δt — приращения переменных x и t , причем $\Delta x = \pi/J$, где J — натуральное число. Множество точек плоскости (x, t) с координатами $x = j\Delta x$, $t = n\Delta t$, $j = 0, 1, 2, \dots, J$ и $n = 0, 1, 2, \dots$, будем называть *сеткой* с размерами ячеек Δx и Δt ; Δx и Δt считаются малыми положительными приращениями, и в дальнейшем мы будем рассматривать предельные процессы, при которых они стремятся к нулю. Приближенное значение величины $u(j\Delta x, n\Delta t)$ будем обозначать через u_j^n . Одной из простейших аппроксимаций дифференциального уравнения (1.2) является конечно-разностное уравнение

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \sigma \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2},$$

$$j = 1, 2, \dots, J-1,$$

$$n = 0, 1, 2, \dots, \quad (1.9)$$

с граничными и начальными условиями

$$u_0^n = 0, \quad u_J^n = 0, \quad n = 0, 1, 2, \dots, \quad (1.10)$$

$$u_j^0 = \varphi(j\Delta x), \quad j = 0, 1, 2, \dots, J, \quad (1.11)$$

которыми заменяются условия (1.4) и (1.3). Ясно, что эти уравнения дают возможность последовательно определить все значения u_j^n , $0 \leq j \leq J$, $n \geq 0$.

Результаты вычислений, выполненных при помощи этих уравнений для рассматриваемой задачи с начальной функцией (1.5), показаны на рис. 1.1 и 1.2. Вычисления выполнялись при $J = 20$ и, следовательно, при $\Delta x = \pi/20$, но при различных, хотя и не очень сильно отличающихся одно от другого значениях Δt , а именно полагалось

$$\frac{\sigma \Delta t}{(\Delta x)^2} = \begin{cases} 5/11 & \text{для рис. 1.1,} \\ 5/9 & \text{для рис. 1.2.} \end{cases}$$

Сплошными линиями изображено точное решение (1.6) с коэффициентами (1.8), а точками — численное решение¹⁾ конечно-

¹⁾ Как и в большинстве иллюстративных примеров этой книги, эти численные значения были получены в процессе учебных занятий на вычислительных машинах УНИВАК и ИБМ 7090 в вычислительном центре Нью-Йорк-

разностных уравнений (1.9)—(1.11). Сравнение результатов показывает, что вычисления с меньшим значением Δt вполне удовлетворительны, тогда как во втором случае погрешность быстро накапливается и после сравнительно малого числа шагов увеличивается до неприемлемой величины. Явление, возникающее во втором случае, называется *неустойчивостью*. Оно не

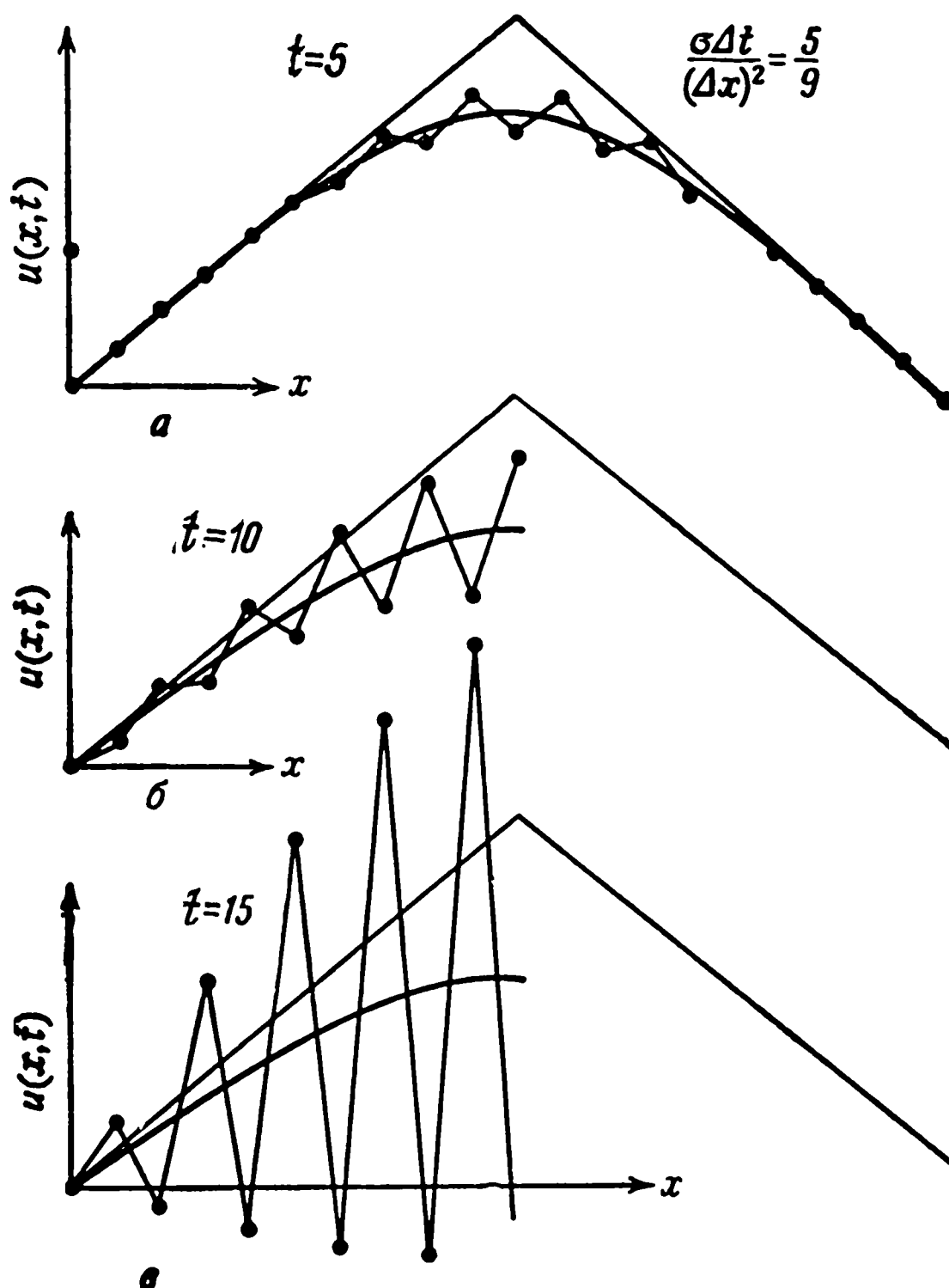


Рис. 1.2. Решение той же задачи, что и на рис. 1.1, полученное при вычислениях с несколько большим значением Δt . Графики а, б и в соответствуют второй, третьей и четвертой (сверху) кривым на рис. 1.1.

связано с ошибками округления, которые в приведенном примере совершенно незначительны, но является внутренним свойством самой системы разностных уравнений (1.9)—(1.11). От-

ского университета. В этом примере, который, конечно, очень прост, задача была запрограммирована так, что ее приближенное и точное решение печатались машиной в виде двух параллельных столбцов для избранных значений n . Точное решение находилось посредством суммирования ряда Фурье с точностью до восьми десятичных знаков,

метим также, что при уменьшении Δx погрешность только увеличивается, если не уменьшать соответственно Δt . Эти факты были отмечены Курантом, Фридрихсом и Леви [1928]¹⁾. Одной из основных целей исследований, представленных в этой книге, является анализ и предупреждение неустойчивости.

§ 1.4. Устойчивость

Если $u(x, t)$ — точное решение краевой задачи, а u_j^n — точное решение конечноразностных уравнений, то *погрешность метода* — это разность $u_j^n - u(j\Delta x, n\Delta t)$. Возникают два вопроса:

1) Как ведет себя $|u_j^n - u(j\Delta x, n\Delta t)|$ при $n \rightarrow \infty$ и фиксированных Δx и Δt ?

2) Как ведет себя $|u_j^n - u(j\Delta x, n\Delta t)|$, когда ячейки сетки измельчаются (т. е. $\Delta x, \Delta t \rightarrow 0$) при фиксированном значении $n\Delta t$?

Мы увидим, что ответ на второй вопрос зависит от отношения скоростей, с которыми Δx и Δt стремятся к нулю. Мы считаем второй вопрос более существенным, потому что основная задача приближенных вычислений — указать способ, посредством которого погрешность можно сделать сколь угодно малой, так чтобы в пределе она стремилась к нулю.

В обоих вопросах число шагов вычислений становится в пределе бесконечным и возможен безграничный рост погрешности. Для однородных краевых задач малость погрешности в обоих случаях обеспечивается совершенно аналогичными условиями на разностные уравнения, но для других краевых задач результаты могут существенно различаться.

Чтобы найти ответы на наши вопросы в случае задачи теплопроводности, заметим, что точное решение разностных уравнений также может быть представлено в виде ряда Фурье. Действительно, пусть A , ξ и m — некоторые константы, причем m — целое число. Подставим в разностное уравнение (1.9) выражение $A\xi^n \exp(imj\Delta x)$ вместо u_j^n . Выбирая ξ так, чтобы уравнение (1.9) при этом удовлетворялось, получим

$$\xi = \xi(m) = 1 - \frac{2\sigma \Delta t}{(\Delta x)^2} (1 - \cos m \Delta x). \quad (1.12)$$

Следовательно, выражение

$$u_j^n = \sum_{m=-\infty}^{\infty} A_m (\xi(m))^n \exp(imj\Delta x), \quad (1.13)$$

где A_m определяются формулой (1.7), является точным решением разностного уравнения (1.9) с граничными условиями

¹⁾ Список оригинальных работ помещен в конце книги; работы обозначаются фамилией их автора с указанием года выхода в свет.

(1.10) и начальными условиями (1.11). В самом деле, (а) ряд (1.13) абсолютно сходится, ибо ряд для $\varphi(x)$ абсолютно сходится и $\xi(m)$ является ограниченной функцией от m ; (б) каждый член ряда (1.13) удовлетворяет разностным уравнениям (1.9), а значит, и его сумма удовлетворяет этим уравнениям; (в) при $n = 0$ ряд (1.13) совпадает с рядом для $\varphi(j\Delta x)$ и, следовательно, удовлетворяет начальным условиям; (г) справедливость граничных условий для функции u_j^n и ее нечетность по j следуют из того, что $A_{-m} = A_m$ в силу нечетности $\varphi(x)$, равенства $\xi(-m) = \xi(m)$ и формулы Эйлера для экспонент. Таким образом, u_j^n удовлетворяет всем уравнениям (1.9) — (1.11).

Отметим между прочим, что так как функция, представленная рядом Фурье (1.13), определяется лишь для последовательности точек $x = j\Delta x$, то в выборе коэффициентов имеется большой произвол. Выбор, сделанный выше, является вполне подходящим, но с таким же успехом u_j^n можно представить и по-другому, полагая например

$$u_j^n = \sum_{m=-J}^J B_m (\xi(m))^n \exp(imj\Delta x),$$

где B_m находятся из начальных условий, рассматриваемых как $2J + 1$ линейных уравнений¹⁾ относительно $2J + 1$ неизвестных $B_{-J}, \dots, B_J$.

Формула (1.13) означает следующее: вследствие того что в рассматриваемой задаче переменные разделяются, каждая гармоника с ростом времени растет (или затухает) независимо от других; множитель перехода $\xi(m)$ является коэффициентом роста (или затухания) амплитуды m -й гармоники для временного интервала Δt . В точном решении (1.6) все гармоники затухают и коэффициент затухания амплитуды m -й гармоники равен $\exp(-m^2\sigma\Delta t)$. Если эту экспоненту и косинус в (1.12) разложить в степенные ряды, то обнаружится, что оба множителя перехода совпадают с точностью до членов, содержащих m^2 :

$$\begin{aligned} \xi(m) &= 1 - m^2\sigma\Delta t + \frac{1}{12} m^4\sigma\Delta t (\Delta x)^2 - \dots, \\ \exp(-m^2\sigma\Delta t) &= 1 - m^2\sigma\Delta t + \frac{1}{2} m^4\sigma^2 (\Delta t)^2 - \dots \end{aligned} \quad (1.14)$$

Для любой гармоники эти два множителя перехода могут быть сделаны сколь угодно близкими посредством соответствующего уменьшения Δx и Δt , и поэтому возникает надежда, что реше-

¹⁾ Начальная функция $\varphi(x)$ продолжена на отрезок $[-\pi, 0]$, так что j теперь изменяется от $-J$ до J .

нием разностного уравнения можно хорошо аппроксимировать решение дифференциального уравнения. С другой стороны, из (1.12) следует, что независимо от малости Δx и Δt всегда имеются некоторые гармоники с достаточно большими номерами m , для которых эти множители значительно различаются. Данное обстоятельство либо лишает законной силы численное решение, либо нет. Ряд Фурье для начальной функции $\varphi(x)$ предполагается абсолютно сходящимся, так что ряд (1.6) для точного решения абсолютно сходится для всех неотрицательных t . Следовательно, когда имеется некоторая надежная степень точности вычислений, то в точном решении все гармоники достаточно высокого порядка, скажем с $m \geq m_0$, при всяком $t \geq 0$ пренебрежимо малы.

Искажение, которое вносят эти гармоники высшего порядка в аппроксимирующее решение, не приносят вреда, если только они не возрастают до величины, которая не является уже пренебрежимо малой. Интуитивно ясно, следовательно, что условием устойчивости является неравенство

$$\max_m |\xi(m)| \leq 1. \quad (1.15)$$

Если это условие выполнено, то ни одна из гармоник не возрастает, если же оно нарушено, то имеются гармоники, неограниченно возрастающие при возрастании n .

Конечно, может случиться, что при специальном выборе начальной функции начальные амплитуды A_m гармоник высших порядков ($m > m_0$) равны нулю. В этом случае они останутся равными нулю и при любой степени усиления в предположении, что арифметические операции при решении разностных уравнений выполняются абсолютно точно (т. е. без всяких округлений). Однако на практике условие (1.15) должно рассматриваться как необходимое.

Мы теперь в состоянии ответить на первый вопрос, поднятый в начале этого параграфа. Погрешность $u_j^n - u(j\Delta x, n\Delta t)$ остается ограниченной при $n \rightarrow \infty$ для фиксированных Δx и Δt и для произвольной начальной функции с абсолютно сходящимся рядом Фурье тогда и только тогда, когда выполняется условие (1.15). Для доказательства заметим сначала, что поскольку точное решение $u(x, t)$ ограничено при $t \rightarrow \infty$, то погрешность метода может быть ограниченной¹⁾ тогда и только тогда, когда u_j^n ограничено. Но

$$|u_j^n| = \left| \sum_{m=-\infty}^{\infty} A_m (\xi(m))^n \exp(im_j \Delta x) \right| \leq \sum_{m=-\infty}^{\infty} |A_m| |(\xi(m))^n|,$$

¹⁾ Мы здесь не делаем никакой попытки доказать, что погрешность метода в этом случае в некотором смысле мала.

так что если выполнено условие (1.15), то

$$|u_j^n| \leq \sum_{m=-\infty}^{\infty} |A_m|,$$

т. е. u_j^n ограничено в силу предположенной в самом начале абсолютной сходимости ряда Фурье для $\varphi(x)$. С другой стороны, если $|\xi(m)| > 1$ для некоторого m , то достаточно взять $\varphi(x) = \sin mx$, чтобы получить неограниченное решение u_j^n , ибо в этом случае u_j^n равно $(\xi(m))^n \sin mj\Delta x$.

Чтобы применить условие (1.15) к задаче теплопроводности, заметим, что в силу формулы (1.12) множитель перехода $\xi(m)$ действителен для всех действительных m и никогда не превосходит $+1$. Условие (1.15) означает, следовательно, что $\xi(m)$ не может быть меньше -1 . Наименьшее значение $\xi(m)$ достигается тогда, когда m является нечетным кратным J . Условие $\xi(m) \geq -1$ для таких значений m состоит в том, что

$$\frac{2\sigma \Delta t}{(\Delta x)^2} \leq 1, \quad (1.16)$$

и это есть условие устойчивости для системы разностных уравнений (1.9) — (1.11).

В рассматриваемом примере достаточность условия (1.16) можно доказать более элементарно. Действительно, замечая, что разностные уравнения (1.9) могут быть переписаны в виде

$$u_j^{n+1} = Au_{j+1}^n + Bu_j^n + Cu_{j-1}^n,$$

где

$$A = C = \frac{\sigma \Delta t}{(\Delta x)^2}, \quad B = 1 - \frac{2\sigma \Delta t}{(\Delta x)^2}$$

(так что если условие (1.16) выполнено, то A , B , C положительны и их сумма равна единице), получаем

$$\begin{aligned} \max_j |u_j^{n+1}| &\leq (A + B + C) \max_j |u_j^n| = \max_j |u_j^n| \leq \\ &\leq \max_j |u_j^{n-1}| \leq \max_j |u_j^0|, \end{aligned}$$

и решение, таким образом, ограничено.

Это доказательство имеет то достоинство, что оно применимо в слегка измененном виде даже в том случае, когда коэффициенты дифференциального уравнения переменны; но метод анализа устойчивости при помощи рядов Фурье, введенный фон Нейманом, может быть применен к более широкому классу дифференциальных уравнений, где элементарные методы теряют силу. Кроме того он дает возможность проникнуть в сущность явлений, происходящих на практике, например иллюстри-

рованных рис. 1.1 и 1.2, где видно, что когда возникает неустойчивость, то те гармоники, длина волны которых сравнима с удвоенным шагом пространственной сетки, неприемлемым образом усиливаются.

Еще один метод, который мы детально рассмотрим в гл. 6, расположен по области применимости где-то между указанными двумя методами. Это так называемый энергетический метод, обязанный своим названием принципу сохранения энергии физической системы, описываемой дифференциальным уравнением. Во избежании возможных недоразумений следует сразу отметить, что во многих случаях сохраняющаяся величина не является физической энергией. Хорошим примером в этом смысле является рассматриваемая задача теплопроводности: здесь физическая энергия пропорциональна величине

$$\int_{-\pi}^{\pi} u \, dx;$$

мы однако рассмотрим вместо нее ¹⁾

$$\int_{-\pi}^{\pi} u^2 \, dx.$$

Умножим обе части уравнения (1.2) на u и проинтегрируем получившееся равенство по x от $-\pi$ до π . Интегрируя правую часть по частям и замечая, что внеинтегральные члены равны нулю в силу граничных условий (1.4), получаем

$$\frac{1}{2} \frac{\partial}{\partial t} \int_{-\pi}^{\pi} u^2 \, dx = -\sigma \int_{-\pi}^{\pi} \left(\frac{\partial u}{\partial x} \right)^2 \, dx \leq 0.$$

Аналогично для разностной схемы (1.9) положим

$$\|u^n\|^2 = \Delta x \sum_{j=1}^{J-1} (u_j^n)^2$$

и постараемся показать, что эта величина остается ограниченной при $n \rightarrow \infty$, если выполнено условие (1.16). Для этого умножим (1.9) на $(u_j^{n+1} + u_j^n) \Delta x$ и просуммируем по j , используя формулу суммирования по частям

$$\begin{aligned} \sum_{j=1}^{J-1} u_j (v_{j+1} - 2v_j + v_{j-1}) &= u_J (v_J - v_{J-1}) - u_1 (v_1 - v_0) - \\ &\quad - \sum_{j=1}^{J-1} (u_{j+1} - u_j) (v_{j+1} - v_j). \end{aligned}$$

¹⁾ Потому что предыдущий интеграл ввиду нечетности u равен нулю. — Прим. перев.

Учитывая еще граничные условия (1.10), находим

$$\begin{aligned} \|u^{n+1}\|^2 - \|u^n\|^2 &= \\ &= -\frac{1}{2} L \left[(u_1^n + u_1^{n+1}) u_1^n \Delta x + \|\Delta_+ u^n\|^2 + \Delta x \sum_{j=1}^{J-1} \Delta_+ u_j^n \Delta_+ u_j^{n+1} \right], \end{aligned} \quad (1.17)$$

где $L = 2\sigma \Delta t / (\Delta x)^2$, а через $\Delta_+ u(x)$ обозначена „разность вперед“ $u(x + \Delta x) - u(x)$. В силу неравенства Шварца

$$\left| \Delta x \sum_{j=1}^{J-1} \Delta_+ u_j^n \Delta_+ u_j^{n+1} \right| \leq \|\Delta_+ u^n\| \|\Delta_+ u^{n+1}\| \leq \frac{1}{2} (\|\Delta_+ u^n\|^2 + \|\Delta_+ u^{n+1}\|^2).$$

Далее, положив

$$S_n = \|u^n\|^2 - \frac{1}{4} L [(u_1^n)^2 \Delta x + \|\Delta_+ u^n\|^2], \quad (1.18)$$

найдем, что

$$\begin{aligned} S_{n+1} - S_n &= -\frac{1}{4} L \left[(u_1^n + u_1^{n+1})^2 \Delta x + \|\Delta_+ u^n\|^2 + \right. \\ &\quad \left. + \|\Delta_+ u^{n+1}\|^2 + 2 \Delta x \sum_{j=1}^{J-1} \Delta_+ u_j^n \Delta_+ u_j^{n+1} \right] \leq 0, \end{aligned} \quad (1.19)$$

т. е. S_n не возрастает с ростом n . Ясно, что $S_0 \leq \|u^0\|^2$, и согласно (1.19) $S_n \leq \|u^0\|^2$. Для завершения доказательства нужно установить, что $\|u^n\|^2 \leq c S_n$, где c — некоторая константа. Так как $u_j^n = 0$, то

$$(u_1^n)^2 + \sum_{j=1}^{J-1} (u_{j+1}^n - u_j^n)^2 \leq 4 \sum_{j=1}^{J-1} (u_j^n)^2.$$

Если $L \leq 1 - \varepsilon$, $\varepsilon > 0$ ¹⁾, то $S_n \geq \|u^n\|^2 - (1 - \varepsilon) \|u^n\|^2 = \varepsilon \|u^n\|^2$, т. е. $\|u^n\|^2$ и, следовательно, u_j^n ограничены при $n \rightarrow \infty$.

Более важным, чем вопрос об ограниченности решения разностных уравнений, является второй вопрос, поставленный в начале этого параграфа, — вопрос о поведении погрешности в данной точке (x, t) , когда ячейки сетки безгранично измельчаются; мы хотим, чтобы при этом погрешность в пределе стремилась к нулю. Предположим, что (x, t) — это точка сетки со сторонами ячеек $\Delta_1 x, \Delta_1 t$, для которой выполняются вычисления по конечно-разностной схеме, и что для улучшения полученного приближения вычисления повторяют, измельчая эту сетку. Значение n , соответствующее фиксированному моменту времени t , при этом,

¹⁾ Здесь приходится исключить случай равенства в (1.16). — Прим. перев.

конечно, неограниченно возрастает, и интуитивно ясно, что нужно избегать случая, когда множитель перехода по абсолютной величине превышает 1, по тем же соображениям, что и выше. Поэтому, рассматривая другую сетку со сторонами ячеек $\Delta x = \Delta_1 x / K$, $\Delta t = \Delta_1 t / K^2$ (K — натуральное число), предположим, что $\Delta_1 x$ и $\Delta_1 t$ таковы, что величина

$$L = \frac{2\sigma \Delta t}{(\Delta x)^2} = \frac{2\sigma \Delta_1 t}{(\Delta_1 x)^2}$$

не превышает 1, так что $|\xi(m)| \leq 1$ для всех m . Пусть $u^n(x, t)$ обозначает приближенное значение $u(x, t)$ в данной точке (x, t) , где $t = n\Delta t$, являющееся решением конечноразностных уравнений. Мы утверждаем, что при высказанных предположениях $u^n(x, t) - u(x, t) \rightarrow 0$ при $K \rightarrow \infty$.

Доказательство. Пусть m_0 — натуральное число (пока произвольное). Для данного K имеем

$$\begin{aligned} |u^n - u| &= \\ &= \left| \left[\sum_{|m| \leq m_0} + \sum_{|m| > m_0} \right] A_m \exp(imx) [\xi^{t/\Delta t} - (\exp(-m^2 \sigma \Delta t))^{t/\Delta t}] \right| = \\ &= |\sum_1 + \sum_2|, \end{aligned} \quad (1.20)$$

где через $\sum_1$ и $\sum_2$ обозначены суммы, соответствующие $|m| \leq m_0$ и $|m| > m_0$. Вторая сумма удовлетворяет неравенству

$$|\sum_2| \leq 2 \sum_{|m| > m_0} |A_m|,$$

так как $|\xi| \leq 1$ и $|\exp(-m^2 \sigma \Delta t)| \leq 1$; следовательно, $\sum_2$ можно сделать сколь угодно малой за счет выбора достаточно большого m_0 , ибо ряд Фурье для $\varphi(x)$ абсолютно сходится.

Для оценки первой суммы заметим, что выражение во второй квадратной скобке в (1.20) имеет вид $\xi^n - \eta^n$, где ξ и η — множители перехода. Оно может быть записано в виде произведения $(\xi - \eta)(\xi^{n-1} + \xi^{n-2}\eta + \dots + \eta^{n-1})$, второй множитель которого не превосходит n , поскольку каждый из n его членов не превосходит 1. Далее, выражение

$$\frac{\xi - \exp(-m^2 \sigma \Delta t)}{m^4 (\Delta t)^2}$$

является аналитической функцией от $m^2 \Delta t$ в некоторой окрестности нуля в силу свойств рядов (1.14) и, очевидно, ограничено для всех действительных $m^2 \Delta t$, больших некоторого фиксированного $\varepsilon > 0$; отсюда следует, что оно ограничено для всех неотрицательных $m^2 \Delta t$. Пусть модуль этого выражения не превы-

шает B . Тогда

$$\left| \sum_1 \right| \leq \sum_{|m| \leq m_0} m^4 (\Delta t)^2 B \frac{t}{\Delta t} |A_m| \leq m_0^4 B t \Delta t \sum_{m=-\infty}^{\infty} |A_m|.$$

Выбрав m_0 столь большим, чтобы $\sum_2$ сделалась достаточно малой, выберем затем Δt столь малым, чтобы сделалась достаточно малой и $\sum_1$. Таким образом, погрешность метода можно сделать сколь угодно малой путем выбора достаточно большого K , и сходимость доказана. Подобное доказательство было предложено Хилдебрандом [1952].

Итак, нами были представлены две точки зрения на устойчивость. Независимо от того, какая из них имеется в виду: t неограниченно возрастает при фиксированном Δt или Δt стремится к нулю при фиксированном t (или, как в данном случае, имеет место и то и другое), необходимо соблюдение условия (1.16), чтобы погрешности (аппроксимации, округления или любого другого рода) не могли достигать величины, при которой всякие вычисления теряют смысл. На практике, конечно, делается только конечное число шагов вычислений с конечной ячейкой, однако признаки неустойчивости появляются уже при сравнительно малом числе шагов, если условие (1.16) не выполнено. Всякий раз, когда для приближенного решения краевой задачи применяются конечноразностные уравнения, следует установить условия, при которых они являются устойчивыми. Такого рода условия обычно представляют собой ограничения, налагаемые на допустимую величину Δt , выраженную через величины других приращений. Однако существуют примеры как безусловно устойчивых, так и безусловно неустойчивых разностных уравнений.

§ 1.5. Неявные разностные уравнения

Условие устойчивости (1.16), полученное для простейшей разностной схемы в предыдущем параграфе, приводит к неблагоприятным выводам: если в интересах повышения точности Δx выбирается достаточно малым, то Δt в свою очередь может оказаться столь малым, что для завершения процесса решения задачи потребуются огромное число шагов по времени, ибо наибольшее допустимое для Δt значение пропорционально квадрату Δx . Ситуация существенно меняется при переходе к системе

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \sigma \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{(\Delta x)^2}, \quad (1.21)$$

которая совпадает с (1.9), за исключением верхних индексов в правой части. Говорят, что системы (1.9) и (1.21) используют разности по времени соответственно *вперед* и *назад* (относи-

тельно того момента времени t , для которого составлены пространственные разности).

Мы увидим, что система (1.21) безусловно устойчива, однако гораздо поучительнее рассмотреть целое семейство разностных систем, правые части которых получаются путем усреднения правых частей уравнений (1.9) и (1.21), взятых с некоторыми весами. Сначала введем некоторые обозначения. Для всякой функции f от x положим

$$(\delta f)_j = f\left(\left(j + \frac{1}{2}\right)\Delta x\right) - f\left(\left(j - \frac{1}{2}\right)\Delta x\right),$$

где j — целое число или целое число плюс $1/2$, так что

$$(\delta^2 f)_j = f((j+1)\Delta x) - 2f(j\Delta x) + f((j-1)\Delta x)$$

и т. д.; если f зависит и от t , то употребляются также и верхние индексы n , $n+1$ и т. д. Используя эти обозначения, напомним систему разностных уравнений, которую мы собираемся рассматривать:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \sigma \frac{\theta (\delta^2 u)_j^{n+1} + (1-\theta) (\delta^2 u)_j^n}{(\Delta x)^2}, \quad (1.22)$$

где θ — действительная постоянная, выбираемая обычно на отрезке $0 \leq \theta \leq 1$.

Если $\theta = 0$, как это было в предыдущем параграфе, то система называется *явной*. В этом случае неизвестное u_j^{n+1} непосредственно находится из соответствующего уравнения по известным величинам u_j^n . Если θ отлично от нуля, то для получения u_j^{n+1} необходимо решить всю систему линейных уравнений (1.22) совместно, в связи с чем ее называют *неявной*.

Подстановка выражения $A\xi^n \exp(imj\Delta x)$ вместо u_j^n в (1.22) показывает, что для данного m множитель перехода имеет вид

$$\xi = \xi(m) = \frac{1 - (1-\theta)L(1 - \cos m\Delta x)}{1 + \theta L(1 - \cos m\Delta x)}, \quad (1.23)$$

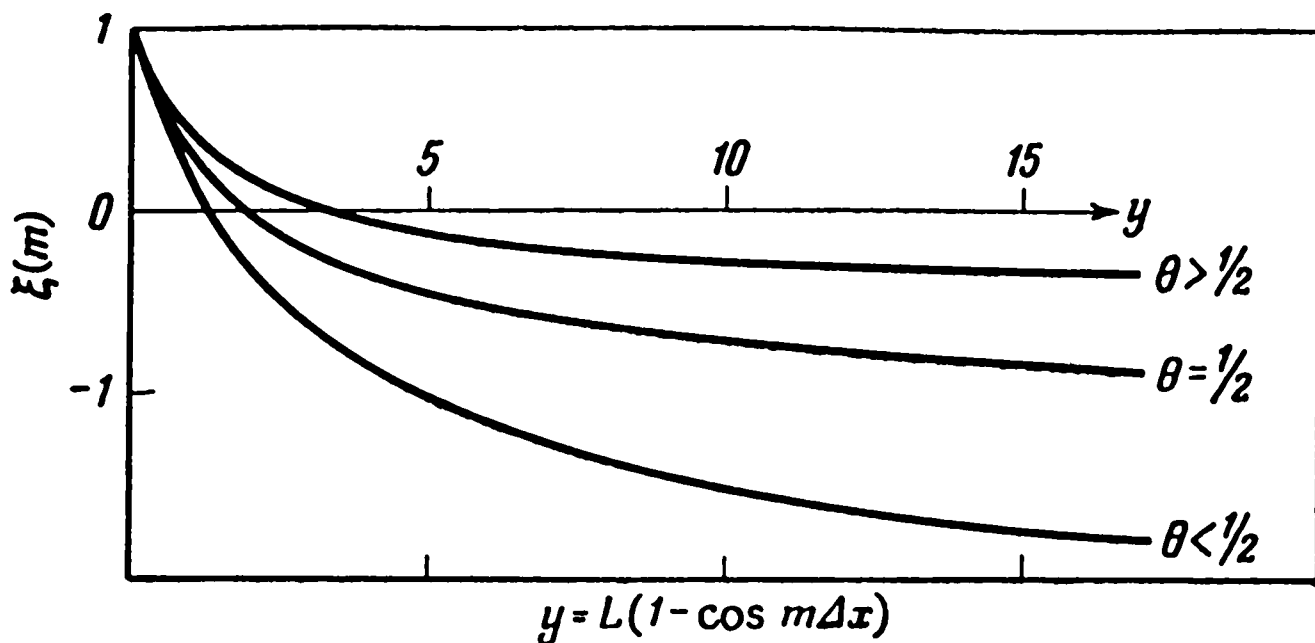
где L , как и прежде, обозначает величину $2\sigma\Delta t/(\Delta x)^2$. И в этом случае имеет силу тот вывод, что функция $\xi(m)$ действительна для всякого действительного m и никогда не превышает $+1$. На рис. 1.3 представлена зависимость величины $\xi(m)$ от аргумента $y = L(1 - \cos m\Delta x)$ для различных значений θ . При возрастании y от 0 до $+\infty$ значение $\xi(m)$ монотонно убывает от 1 до $-(1-\theta)/\theta$. Если $1/2 \leq \theta \leq 1$, то это предельное значение лежит не ниже прямой $\xi = -1$ и, следовательно, система (1.22) безусловно устойчива. Но если $0 \leq \theta < 1/2$, то для обеспечения устойчивости нужно ограничить y тем его значением, при котором график функции $\xi(m)$ пересекается с прямой $\xi = -1$.

В соответствии с этим условие устойчивости формулируется так:

$$\frac{2\sigma \Delta t}{(\Delta x)^2} \leq \frac{1}{1-2\theta} \quad \text{при} \quad 0 \leq \theta < \frac{1}{2}, \quad (1.24)$$

никаких ограничений при $\frac{1}{2} \leq \theta \leq 1$.

Достаточность этого условия для устойчивости можно без труда установить также и энергетическим методом. Преобразуя



Р и с. 1.3. Множитель перехода для неявного разностного уравнения (1.22); см. также формулу (1.23).

(1.22) аналогично тому, как это было сделано для (1.9), мы получим для $\|u^{n+1}\|^2 - \|u^n\|^2$ выражение, сходное с (1.17). Далее, для величины

$$S_n = \|u^n\|^2 - \frac{1}{4} (1 - 2\theta) L [(u_1^n)^2 \Delta x + \|\Delta_+ u^n\|^2] \quad (1.25)$$

получается такое же условие (1.19), как и при $\theta = 0$. Следовательно, $\|u^n\|^2 \leq \text{const } S_n \leq \text{const } S_0 \leq \text{const } \|u^0\|^2$, если только выполнено условие $(1 - 2\theta)L < 1$, которое, за исключением предельного случая точного равенства, совпадает с (1.24).

Наиболее ходовыми значениями θ являются 0, 1/2, 1. Первое значение дает явное уравнение (1.9), другие два дают неявные уравнения, которые безусловно устойчивы. При $\theta = 1$ получается уравнение (1.21). Используя неявные схемы, приводящие к решению совместных систем уравнений (практически это оказывается не слишком сложным, см. гл. 8), можно обойти все затруднения, связанные с устойчивостью, и выбирать значения Δt исключительно из соображений желаемой точности.

В § 8.6 приводится численный пример, в котором разностная схема (1.22) применяется к нелинейному уравнению

$$\frac{\partial u}{\partial t} = \frac{\partial^2 (u^5)}{\partial x^2}.$$

Здесь переменная величина $5u^4$ играет роль σ . Результаты показывают, что условие (1.24) при замене σ на $5u^4$ по-прежнему обеспечивает устойчивость, как и следовало ожидать.

§ 1.6. Погрешность аппроксимации

Согласно сказанному в § 1.4 решение явной разностной системы (1.9)—(1.11) сходится к точному решению краевой задачи, если Δx и Δt стремятся к нулю так, что выполняется условие устойчивости (1.16). Соответствующее утверждение справедливо и для неявных систем. Для практических целей определение скорости сходимости также имеет большой интерес; в этом параграфе вопрос о скорости сходимости будет рассмотрен в связи с вопросом о погрешности аппроксимации.

Пусть $\bar{u}(x, t)$ — точное решение уравнения (1.2) с непрерывными частными производными всех встречающихся ниже порядков. Обозначим $\bar{u}(j\Delta x, n\Delta t)$ через $\bar{u}_j^n$, допуская, что j и n могут принимать и нецелые значения. Запишем разложения в ряд Тэйлора с остаточным членом в форме Лагранжа:

$$\begin{aligned}\bar{u}_j^{n+1} &= \bar{u}_j^n + \left(\frac{\partial \bar{u}}{\partial t}\right)_j^n \Delta t + \frac{1}{2} \left(\frac{\partial^2 \bar{u}}{\partial t^2}\right)_j^{n+\theta_1} (\Delta t)^2, \\ \bar{u}_{j+1}^n &= \bar{u}_j^n + \left(\frac{\partial \bar{u}}{\partial x}\right)_j^n \Delta x + \frac{1}{2} \left(\frac{\partial^2 \bar{u}}{\partial x^2}\right)_j^n (\Delta x)^2 + \frac{1}{6} \left(\frac{\partial^3 \bar{u}}{\partial x^3}\right)_j^n (\Delta x)^3 + \\ &\quad + \frac{1}{24} \left(\frac{\partial^4 \bar{u}}{\partial x^4}\right)_{j+\theta_2}^n (\Delta x)^4;\end{aligned}$$

аналогичное разложение имеет место и для $\bar{u}_{j-1}^n$; θ_1 и θ_2 заключены между 0 и 1. Так как $\partial \bar{u} / \partial t = \sigma \partial^2 \bar{u} / \partial x^2$, то

$$\begin{aligned}\frac{\bar{u}_j^{n+1} - \bar{u}_j^n}{\Delta t} - \sigma \frac{\bar{u}_{j+1}^n - 2\bar{u}_j^n + \bar{u}_{j-1}^n}{\Delta x^2} &= \\ &= \frac{1}{2} \left(\frac{\partial^2 \bar{u}}{\partial t^2}\right)_j^{n+\theta_1} \Delta t - \frac{\sigma}{24} \left[\left(\frac{\partial^4 \bar{u}}{\partial x^4}\right)_{j+\theta_2}^n + \frac{\sigma}{24} \left(\frac{\partial^4 \bar{u}}{\partial x^4}\right)_{j+\theta_3}^n \right] (\Delta x)^2.\end{aligned}$$

Коэффициенты при Δt и $(\Delta x)^2$ в правой части этого равенства ограничены в силу непрерывности соответствующих частных производных, так что

$$\frac{\bar{u}_j^{n+1} - \bar{u}_j^n}{\Delta t} - \sigma \frac{(\delta^2 \bar{u})_j^n}{(\Delta x)^2} = O(\Delta t) + O((\Delta x)^2) \quad \text{при } \Delta t, \Delta x \rightarrow 0. \quad (1.26)$$

Отсюда видно, что существуют такие две зависящие от $\bar{u}$ положительные постоянные K_1 и K_2 , что абсолютная величина левой части не превосходит $K_1 \Delta t + K_2 (\Delta x)^2$ для всех достаточно малых Δt и Δx . В тех случаях, когда не может возникнуть недоразумения, слова «при $\Delta t, \Delta x \rightarrow 0$ » опускаются.

Выражение в левой части равенства (1.26) мы будем называть *погрешностью аппроксимации* разностной схемы (1.9), хотя некоторые авторы называют его так только после умножения на Δt .

Равенство (1.26) показывает, насколько хорошо решение дифференциального уравнения удовлетворяет разностным уравнениям при безграничном измельчении ячеек сетки. С его помощью удастся иногда определить и скорость сходимости решения разностной системы к решению дифференциального уравнения. Особенно просто это делается для явной схемы. А именно, если мы обозначим через e_j^n разность $u_j^n - \bar{u}_j^n$, то после вычитания (1.26) из (1.9) получим

$$e_j^{n+1} = \frac{1}{2} L e_{j-1}^n + (1 - L) e_j^n + \frac{1}{2} L e_{j+1}^n + O((\Delta t)^2) + O((\Delta x)^2 \Delta t).$$

Следовательно, если выполнено условие устойчивости $L \leq 1$, то

$$\max_j |e_j^{n+1}| \leq \max_j |e_j^n| + K_1 (\Delta t)^2 + K_2 (\Delta x)^2 \Delta t.$$

Поэтому при $n\Delta t \leq t$

$$\max_{n, j} |e_j^n| \leq t [K_1 \Delta t + K_2 (\Delta x)^2].$$

т. е. $|e_j^n|$ всего лишь в t раз может превысить оценку для погрешности аппроксимации на одном шаге. Отсюда при фиксированном L и $\Delta x \rightarrow 0$ и следует сходимость.

Включая в (1.26) члены более высокого порядка малости, чем Δt и $(\Delta x)^2$, и учитывая, что функция $\bar{u}$ как решение уравнения теплопроводности удовлетворяет также уравнению

$$\frac{\partial^2 \bar{u}}{\partial t^2} = \sigma^2 \frac{\partial^4 \bar{u}}{\partial x^4},$$

получаем для главного члена погрешности аппроксимации выражение

$$\left(\frac{\sigma^2 \Delta t}{2} - \frac{\sigma (\Delta x)^2}{12} \right) \frac{\partial^4 \bar{u}}{\partial x^4}.$$

Если предположить, что Δt и Δx стремятся к нулю таким образом, что

$$\frac{\sigma \Delta t}{(\Delta x)^2} = \frac{1}{6},$$

то этот главный член обращается в нуль и погрешность аппроксимации в данном случае¹⁾ не превосходит $O((\Delta t)^2)$, или, что то же самое, $O((\Delta x)^4)$. В этом частном случае погрешность ап-

¹⁾ Заметим также, что в этом случае разложения (1.14) совпадают с точностью до членов, содержащих m^4 .

проксимации стремится к нулю даже быстрее, чем утверждает равенство (1.26), и можно ожидать, что решение разностного уравнения приближается к решению дифференциального уравнения особенно быстро. Это действительно имеет место, и некоторые авторы воспользовались данным обстоятельством для повышения точности численного решения простейшей задачи теплопроводности. К сожалению, пока не существует достаточно простого обобщения этого приема на общую задачу теплопроводности или диффузии с переменными коэффициентами. Однако можно построить некоторые другие обобщенные схемы повышенной точности. Две такие схемы для задачи теплопроводности даны в § 8.3.

Вообще говоря, чем меньше погрешность аппроксимации, тем быстрее численное решение сходится к точному. Но следует заметить, что соотношение (1.26) остается в силе и тогда, когда условие устойчивости нарушается; фактически сущность неустойчивости можно охарактеризовать в следующей на первый взгляд парадоксальной форме: если ячейки сетки безгранично измельчаются при нарушении условия устойчивости, то точное решение дифференциального уравнения все лучше и лучше удовлетворяет разностным уравнениям, в то время как точное решение разностных уравнений в общем случае все больше и больше отклоняется от точного решения дифференциального уравнения.

§ 1.7. Скорость сходимости

Оценку скорости сходимости для разностного уравнения (1.9) можно получить с помощью изучения вспомогательных сумм Σ_1 и Σ_2 , на которые была разбита погрешность метода при доказательстве сходимости в § 1.4 (см. равенство (1.20)). Погрешность метода была определена как разность между решением u_j^n разностного уравнения (1.9) и решением u дифференциального уравнения (1.2) в некоторой фиксированной точке (x, t) , принадлежащей бесконечной последовательности сеток с безгранично уменьшающимися ячейками; мы хотим более подробно, чем в предыдущем параграфе, исследовать скорость убывания этой погрешности при соответствующей последовательности вычислений. Покажем, что результат зависит от асимптотического поведения коэффициентов Фурье начальной функции $\varphi(x)$, которое в свою очередь определяется гладкостью этой функции, в соответствии со следующей леммой¹⁾: *если функ-*

¹⁾ См. Зигмунд [1952], где приведено доказательство для случая $p = 0$; для $p > 0$ результат получается p -кратным интегрированием по частям.

ция $\varphi(x)$ и ее производные до $(p-1)$ -го порядка включительно непрерывны, а p -я производная имеет ограниченную вариацию, то асимптотическое поведение коэффициентов Фурье для $\varphi(x)$ характеризуется равенством

$$A_m = O\left(\frac{1}{m^{p+1}}\right) \quad \text{при } m \rightarrow \infty. \quad (1.27)$$

Ограниченность Σ_2 сразу следует из этой леммы: при $p > 0$

$$\left| \Sigma_2 \right| \leq 2 \sum_{|m| > m_0} |A_m| = O\left(\frac{1}{m_0^p}\right). \quad (1.28)$$

Применяя (1.27) к полученной в § 1.4 оценке для Σ_1 , находим

$$\left| \Sigma_1 \right| \leq \sum_{|m| \leq m_0} m^4 (\Delta t)^2 B \frac{t}{\Delta t} |A_m| \leq c_0 B t \Delta t \sum_{1 \leq |m| \leq m_0} |m|^{4-p-1}.$$

При $p \leq 3$ последнее слагаемое в этой сумме является наибольшим и по нему можно оценить всю сумму; при $p = 4$ сумма возрастает с ростом m_0 лишь логарифмически, а при $p > 4$ она ограничена. Поэтому

$$\left| \Sigma_1 \right| = \begin{cases} C_1 m_0^{4-p} \Delta t, & \text{если } p \leq 3, \\ C_2 (\ln m_0) \Delta t, & \text{если } p = 4, \\ C_3 \Delta t, & \text{если } p > 4, \end{cases}$$

где C_1, C_2, C_3 — константы, зависящие от x, t, p, σ и т. д., но не зависящие от Δt и m_0 . Выберем теперь m_0 таким образом, чтобы минимизировать оценку для $|\Sigma_1| + |\Sigma_2|$. Как оказывается, при $p \leq 4$ такое m_0 должно быть пропорционально величине $(\Delta t)^{-1/4}$, так что окончательная оценка для погрешности метода имеет следующий вид:

$$u^n - u = \begin{cases} O((\Delta t)^{p/4}) = O((\Delta x)^{p/2}), & \text{если } p \leq 3, \\ O(\Delta t \ln \Delta t) = O((\Delta x)^2 \ln \Delta x), & \text{если } p = 4, \\ O(\Delta t) = O((\Delta x)^2), & \text{если } p > 4. \end{cases} \quad (1.29)$$

Если начальная функция $\varphi(x)$ кусочно линейна, как например функция (1.5), то $p = 1$ и погрешность метода стремится к нулю как $(\Delta x)^{1/2}$ при $\Delta x \rightarrow 0$; в другом крайнем случае, когда $\varphi(x)$ является аналитической функцией, погрешность убывает как $(\Delta x)^2$.

Если, как и раньше, $\Delta t, \Delta x$ стремятся к нулю таким образом, что $\sigma \Delta t / (\Delta x)^2 = 1/6$, то

$$u^n - u = \begin{cases} O((\Delta t)^{p/3}), & \text{если } p \leq 5, \\ O((\Delta t)^2 \ln \Delta t), & \text{если } p = 6, \\ O((\Delta t)^2), & \text{если } p > 6. \end{cases}$$

Опуская доказательство, которое в этом случае усложняется, отметим лишь, что для гладких начальных данных ($p > 6$) более точное разностное уравнение дает соответственно более высокую скорость сходимости, в то время как для негладких данных ($p = 1$) скорость сходимости увеличивается лишь незначительно ($(\Delta t)^{1/3}$ вместо $(\Delta t)^{1/4}$), причем даже такое незначительное увеличение возможно только потому, что решение уравнения теплопроводности с любыми начальными данными является аналитической функцией при всех $t > 0$. Для гиперболических задач с негладкими начальными данными применение более точных разностных уравнений вообще не дает улучшения сходимости.

Если предположить, что точное решение $u(x, t)$ нашей задачи имеет непрерывные производные по x до четвертого порядка включительно, то, как показывает полученная в § 1.6 более простая оценка для погрешности метода, погрешность при этом стремится к нулю не медленнее, чем Δt .

§ 1.8. Замечания о формулах высших порядков и погрешностях округления

Хорошо известно, что при решении дифференциальных уравнений в частных производных конечноразностные формулы, обеспечивающие высокий порядок точности, например имеющие погрешность аппроксимации порядка $O((\Delta t)^4)$, обычно весьма неудобны на практике. Противоположная картина имеет место для обыкновенных дифференциальных уравнений, для которых такие методы, как метод Рунге — Кутта, часто обеспечивают высокий порядок точности при малой затрате труда; эти методы не имеют аналогов для уравнений в частных производных. Причина этого различия лежит в самой сущности вещей. Для обыкновенного дифференциального уравнения с независимой переменной t задания конечного (и обычно совсем небольшого) числа величин в момент времени $t = t_0$ достаточно (по крайней мере в принципе) для полного и точного определения решения в любой момент $t > t_0$, и точность приближенного решения для $t = t_0 + \Delta t$ зависит только от искусства, с которым используется при вычислениях имеющаяся информация. Для уравнения в частных производных, в котором одной из независимых переменных является t , необходимо задать при $t = t_0$ значения бесконечного числа величин (или какой-нибудь их эквивалент), чтобы определить решение для $t > t_0$, и потому точность приближенного решения для $t = t_0 + \Delta t$ ограничивается не только возможными недостатками метода вычислений, но и недостатком нужной информации.

Если помимо самих значений начальных функций в точках сетки известно, скажем, что они непрерывно дифференцируемы определенное число раз, то это будет дополнительной информацией. Поэтому для задач с достаточно гладкими начальными функциями можно несколько увеличить точность расчета за счет применения более точных разностных формул, при том однако условии, что все граничные значения, особенности, скачки на внутренних границах и т. д. будут аппроксимироваться с соответствующей точностью. Мы вернемся к этому вопросу в гл. 13.

В любом случае, однако, при решении уравнений в частных производных конечноразностными методами обычно достигается более чем скромная точность. По мнению авторов погрешности округления при этом вообще не имеют большого значения. При использовании устойчивой разностной схемы погрешности округления не усиливаются с течением времени; они просто накапливаются примерно пропорционально корню квадратному из числа шагов вычислений, если округление случайно и если заботиться о том, чтобы значения рассматриваемых величин не выходили за определенные границы. Конечно, современные быстродействующие автоматические вычислительные машины способны производить вычисления с очень большим числом шагов, что часто и делается, так что ошибки округления могут заметно накапливаться. Но эти машины обычно имеют разрядную сетку с достаточно большим числом десятичных или двоичных разрядов, и увеличение быстродействия машин примерно в миллион раз может быть компенсировано, с точки зрения накопления случайных погрешностей округления, просто введением трех дополнительных десятичных знаков. Кроме того, предусмотренное на многих машинах выполнение арифметических операций с переменным числом значащих цифр (см. Эшенхорст и Метрополис, 1959) позволяет регулировать влияние погрешностей округления более или менее автоматически, тогда как влияние погрешностей аппроксимации можно оценить только теоретически.

§ 1.9. Содержание следующих глав

Как видно из предыдущего, основные вопросы при использовании разностных методов типа описанных в этой главе это: (1) сходимость к точному решению при безграничном измельчении сетки (в рассмотренном примере все зависело от устойчивости и от порядка малости погрешности аппроксимации); (2) скорость сходимости; (3) построение устойчивых разностных схем с быстрой сходимостью; (4) методы решения неявных разностных систем. Для более сложных задач возникают дополнительные вопросы: (5) влияние переменных коэффициентов и нелинейностей на устойчивость; (6) аппроксимация внешних

и внутренних граничных условий и их влияние на устойчивость; (7) отыскание практических критериев устойчивости в тех случаях, для которых формальная теория либо слишком сложна, либо вообще неприменима. В первой части этой книги вопросы (1)—(6) рассматриваются с теоретической точки зрения, во второй части рассмотрены с точки зрения практики различные приложения теории к решению краевых задач математической физики.

Общая теория для линейных задач изложена во второй и третьей главах. Основные понятия определяются в терминах теории линейных операторов в банаховом пространстве и приводят к теореме Лакса об эквивалентности устойчивости и сходимости. А именно, вводится понятие корректно поставленной краевой задачи и рассматривается конечноразностная аппроксимация для нее; формулируется условие согласованности, которое служит для проверки того, что разностная система аппроксимирует именно данную, а не какую-либо другую краевую задачу; после этого устойчивость аппроксимации определяется как равномерная ограниченность некоторого множества операторов. Теорема Лакса утверждает тогда, что если для корректно поставленной краевой задачи выполнено условие согласованности, то устойчивость является необходимым и достаточным условием сходимости решения разностных уравнений к точному решению краевой задачи при произвольных начальных данных. Другие возможные определения устойчивости кратко рассмотрены в § 5.2 (см. также § 3.9).

В гл. 4 эта теория применяется к линейным задачам с постоянными коэффициентами и однородными граничными условиями. Как было отмечено много лет назад еще фон Нейманом (см. О'Брайен, Хайман, Каплан [1951]), в этом случае исследование можно провести при помощи преобразования Фурье в особенно простой и удобной форме, которая позволяет сразу получить необходимое условие устойчивости фон Неймана. После выхода в свет первого издания этой книги в работах Като, Баченана и Крайса были получены общие достаточные условия устойчивости для таких задач; эти результаты излагаются подробно с упрощенными по возможности доказательствами.

Глава 5 посвящена задачам с переменными коэффициентами (в основном параболическим и гиперболическим); в последнем параграфе кратко рассмотрены некоторые нелинейные (квазилинейные) задачи. Вводится важное понятие диссипативной разностной схемы, принадлежащее Крайсу. Для таких схем полностью оправдала себя идея фон Неймана рассматривать устойчивость локально. По-видимому, на практике можно всегда (без всякой потери точности) пользоваться диссипативными схемами для решения гиперболических и других задач,

дифференциальная форма когорых не является диссипативной. Наконец, в последнем параграфе показано, на основании опытных данных, что введение диссипативных членов надлежащего вида часто является хорошим средством против неустойчивости в нелинейных задачах.

В гл. 6 рассматривается вопрос о влиянии граничных условий на устойчивость в основном для задач с одной пространственной переменной. При этом используются оба основных метода для исследования устойчивости в задачах такого рода: энергетический метод и спектральный метод Годунова и Рябенского; эти методы часто дополняют друг друга.

Глава 7 является своего рода дополнением к первой части книги. В ней результаты гл. 3 для линейных задач обобщаются на случай многослойных разностных систем.

Вторая часть посвящена конкретным практическим особенностям использования методов, уже нашедших широкое применение при решении прикладных задач или находящихся еще в стадии экспериментальной проверки (примером последних являются методы многомерной гидродинамики). При этом иногда полученные в первой части теоретические условия устойчивости, характеризующие поведение вычислительного процесса в случае стремления к нулю приращений аргументов, приходится заменять практическими критериями устойчивости, учитывающими конечность этих приращений (см., например, § 10.5).

В заключение следует отметить, что изложенные в этой книге исследования тесно примыкают к целому ряду превосходных исследований советских математиков в этой области. Например, Ладыженская [1952] доказала устойчивость и сходимост неясной схемы типа Кранка — Николсона для общих гиперболических систем. Общим вопросам конечноразностных методов посвящены работа Ладыженской [1957а] и книга Годунова и Рябенского [1962]. Разрывные решения гиперболических уравнений, с которыми мы будем иметь дело в гл. 13, были рассмотрены в работах Ладыженской [1957б], Олейник и Введенской [1957], Рождественского [1960] и Яненко [1964].

Глава 2

ЛИНЕЙНЫЕ ОПЕРАТОРЫ

§ 2.1. Краевые задачи и функциональные пространства

При решении краевых задач временная переменная t играет особую роль. На любой стадии процесса приходится иметь дело с одной или несколькими функциями некоторых других переменных, которые мы назовем пространственными переменными; эти функции описывают мгновенное состояние физической системы. В случае применения вычислительных машин затабулированные значения этих функций хранятся в ячейках памяти машины. С течением времени состояние физической системы меняется в соответствии с дифференциальными уравнениями, и функции принимают новые значения. Удобно истолковывать эти функции для фиксированного t как точки функционального пространства и обозначать их одним символом u . При такой интерпретации состояние физической системы изображается точкой функционального пространства, а изменение ее состояния во времени представляется движением изображающей точки в этом пространстве.

Для изучения свойств приближений и погрешностей необходимо ввести меру различия между двумя состояниями физической системы; эту меру можно интерпретировать как расстояние между двумя представляющими эти состояния точками.

Если два состояния физической системы почти одинаковы, то расстояние между изображающими их точками должно быть очень малым, и наоборот. Например, если u — непрерывная функция некоторых пространственных переменных, скажем температура в задаче теплопроводности, то в качестве меры различия между двумя состояниями системы можно взять максимум модуля разности соответствующих функций $u_1(x)$ и $u_2(x)$, т. е. максимум функции $|u_1(x) - u_2(x)|$. Другой часто используемый выбор меры различия — это среднеквадратичное отклонение $[(1/V) \int |u_1(x) - u_2(x)|^2 dv]^{1/2}$, где интеграл берется по некоторой области объема V .

В настоящей главе и следующих пяти главах мы ограничимся рассмотрением линейных задач. В этом случае за меру различия двух состояний u и v всегда можно принять некоторую характеристику разности $w = u - v$, которую можно истолковать как абсолютную величину элемента w . Эта величина

называется нормой элемента w и обозначается символом $\|w\|$. Ясно, что норма w должна быть положительным числом, если u и v изображают различные состояния системы, и должна быть равна нулю, если они изображают одинаковые ее состояния, потому что норма представляет собой как раз меру различия состояний. Интерпретация меры различия как расстояния подсказывает в качестве следующего требования выполнение неравенства треугольника, т. е. требование, чтобы норма суммы двух элементов $u + v$ не превышала суммы норм этих элементов u и v , в соответствии с тем геометрическим фактом, что длина стороны* треугольника не превышает суммы длин двух других его сторон. Может показаться не очевидным, что это требование диктуется физическими соображениями, но в действительности неравенство треугольника удовлетворяется при любом выборе нормы, возможном на практике. Предположение о том, что это неравенство выполняется, упрощает теорию операторов, так же как и некоторые другие предположения, которые будут формально введены в следующем параграфе.

Если, как и прежде, u представляет собой некоторую непрерывную функцию, норма которой определяется как максимум ее модуля, то сходимости¹⁾ последовательности точек в функциональном пространстве отвечает равномерная сходимость соответствующих функций. С другой стороны, если u представляет собой измеримую интегрируемую с квадратом функцию, норма которой определяется как среднеквадратичное значение, то сходимости последовательности точек отвечает сходимость в среднем²⁾ соответствующих функций.

Под суммой и разностью двух элементов u и v подразумевают, конечно, функцию (или систему функций), получаемую при сложении или вычитании функций (или систем функций), обозначенных через u и v , причем предполагается, что сумма и разность u и v снова являются элементами функционального пространства. При этом мы можем получить элементы, которые не интерпретируются непосредственно как состояния физической системы. Например, если задачу, разобранную в первой главе, рассматривать как задачу о диффузии газа через некоторую проницаемую среду, то отрицательные значения функции будут лишены реального смысла, так как число молекул не может быть отрицательным. Но несмотря на это целесообразно включать такие функции в функциональное пространство, считая, что они представляют в некотором смысле обобщенные состояния системы; тогда, например, отдельные члены и частные

¹⁾ Под которой понимается тот факт, что расстояние между переменной точкой и предельной точкой стремится к нулю.

²⁾ См. Курант и Гильберт [1951], т. I, стр. 50.

суммы ряда Фурье будут представляться точками функционального пространства. Удобно пойти еще дальше и включить в функциональное пространство комплексные функции в соответствии с определенными правилами, характеризующими линейное пространство.

В силу этих соображений ясно, что аксиомы теории банаховых пространств как раз пригодны для характеристики функциональных пространств, возникающих при решении линейных краевых задач. Необходимые сведения из теории линейных операторов в банаховых пространствах будут изложены в следующих параграфах настоящей главы.

Значения функции, определяемые в результате вычислений по конечноразностным формулам, не дают, конечно, полного представления этой функции. Но тем не менее такие совокупности значений функции могут быть представлены определенными точками функционального пространства, если указан какой-либо линейный закон или правило, при помощи которого можно найти значения функции в точках, не совпадающих с точками сетки.

Например, можно определить промежуточные значения функции с помощью линейной интерполяции. Другой простой прием, полезный при решении краевых задач типа рассматриваемых в гл. 4, это допустить, что разностные уравнения сохраняют силу при всех значениях пространственных переменных. Так, при решении задачи о теплопроводности уравнение (1.9) при этом нужно записать в виде

$$\frac{u^{n+1}(x) - u^n(x)}{\Delta t} = \sigma \frac{u^n(x + \Delta x) - 2u^n(x) + u^n(x - \Delta x)}{(\Delta x)^2};$$

тогда если начальная функция $\varphi(x)$ задана для всех значений x , то из этого уравнения $u^n(x)$ последовательно ($n = 0, 1, 2, \dots$) определяется для всех x ; в действительности, конечно, $u^n(x)$ вычисляется только для конечного множества значений x .

Некоторые авторы рассматривают и точные, и приближенные функции как сеточные, т. е. определенные только в точках сетки, а норму для них обычно определяют как максимум модуля функции. На наш взгляд, это в меньшей степени соответствует физическим представлениям о численных значениях как о точках, через которые проходят кривые или поверхности, приближенно описывающие поведение функции во всей рассматриваемой области пространства.

§ 2.2. Банаховы пространства

Функциональные пространства, используемые в дальнейшем при решении линейных краевых задач, являются по предположению абстрактными пространствами Банаха [1932].

Банахово пространство $\mathcal{B}$ — это, во-первых, *линейное* (векторное) пространство; это значит, что оно состоит из элементов, для которых определены действия сложения и умножения на число (действительное или комплексное); эти действия подчинены следующим правилам.

Если u, v, w — элементы (называемые также точками) пространства $\mathcal{B}$, а a, b, c — некоторые числа, то

$u + v$ — элемент пространства $\mathcal{B}$, au — элемент пространства $\mathcal{B}$,

$$u + v = v + u, \quad a(bu) = (ab)u,$$

$$(a + b)u = au + bu,$$

$$u + (v + w) = (u + v) + w, \quad a(u + v) = au + av^1).$$

Очевидно, что функциональные пространства, встречающиеся при решении линейных краевых задач, удовлетворяют этим аксиомам, если, как указывалось выше, сложение двух элементов определять естественным образом как сложение соответствующих значений функций, а умножение на a — как умножение значений функции на a .

Во-вторых, $\mathcal{B}$ — *нормированное* пространство; это значит, что каждому элементу u пространства $\mathcal{B}$ сопоставлено конечное неотрицательное число, обозначаемое через $\|u\|$, такое, что

$$\|au\| = |a| \|u\|,$$

$$\|u + v\| \leq \|u\| + \|v\|,$$

$$\|u - v\| = 0 \text{ тогда и только тогда, когда } u = v$$

(„ $u - v$ “ — сокращенное обозначение для „ $u + (-1)v$ “).

Наконец, $\mathcal{B}$ — *полное* пространство; это значит, что если последовательность $u_1, u_2, \dots$ элементов из $\mathcal{B}$ такова, что $\|u_n - u_m\| \rightarrow 0$, когда $n, m \rightarrow \infty$ независимо друг от друга (такие последовательности называются *фундаментальными*), то в $\mathcal{B}$ существует элемент u , называемый пределом этой последовательности, такой, что $\|u_n - u\| \rightarrow 0$ при $n \rightarrow \infty$.

Обычным примером функционального банахова пространства является следующий: пусть $\mathcal{B}$ — класс непрерывных функций $f(x)$, определенных на отрезке $a \leq x \leq b$, и пусть норма элемента $f(x)$ определяется как максимум $|f(x)|$ на этом отрезке. Легко проверить, что все перечисленные выше аксиомы выполняются. Сходимость в $\mathcal{B}$ соответствует равномерной сходимости функций, а полнота $\mathcal{B}$ устанавливается следующим об-

¹⁾ В определении линейного пространства надо, кроме того, требовать, чтобы существовал нулевой элемент и чтобы для каждого элемента существовал обратный (противоположный), а также чтобы выполнялось условие $1 \cdot u = u$. — Прим. ред.

разом: последовательность $f_n(x)$, фундаментальная в $\mathcal{B}$, сходится в каждой точке x отрезка $[a, b]$ в силу критерия Коши, причем предельная функция является элементом $\mathcal{B}$, ибо предел равномерно сходящейся последовательности непрерывных функций есть непрерывная функция. Это функциональное пространство и его обобщения очень часто встречаются в численном анализе, но для целей, поставленных в этой книге, более подходящим является пространство с гильбертовой нормой

$$\left[\int_a^b |f(x)|^2 dx \right]^{1/2}$$

(см. гл. 4).

Теперь мы сформулируем некоторые простые следствия из приведенных выше аксиом (читателю, интересующемуся аксиоматическим подходом, стоит вывести их из аксиом¹⁾), однако для целей этой книги достаточно заметить, что они верны в линейных функциональных пространствах). Эти следствия таковы:

$$1 \cdot u = u,$$

$$u - u = v - v = w - w \quad \text{и т. д.}$$

(тем самым определяется элемент $\emptyset = u - u$, называемый нулевым элементом пространства $\mathcal{B}$),

$$\|\emptyset\| = 0, \quad 0 \cdot u = \emptyset, \quad a\emptyset = \emptyset,$$

$$\|u - v\| \geq |\|u\| - \|v\||.$$

Если последовательность элементов пространства $\mathcal{B}$ имеет предел, то этот предел единствен.

Определим некоторые понятия, относящиеся к множествам элементов пространства $\mathcal{B}$. Элемент u называется *предельной точкой* множества $\mathcal{P}$, если сколь угодно близко от точки u имеются точки множества $\mathcal{P}$, т. е. если u является пределом некоторой последовательности, содержащейся в $\mathcal{P}$.

Множество элементов называется *замкнутым*, если оно содержит все свои предельные точки. Замыканием данного множества называется множество, получающееся присоединением к нему (если оно уже не является замкнутым) всех его предельных точек. Множество $\mathcal{P}$ называется *плотным* в $\mathcal{B}$, если произвольный элемент пространства $\mathcal{B}$ можно аппроксимировать сколь угодно точно элементами множества $\mathcal{P}$, т. е. если замыкание $\mathcal{P}$ совпадает с $\mathcal{B}$. Например, если $\mathcal{B}$ — класс непрерывных функций от x с нормой максимум модуля, то множество полиномов от x плотно в $\mathcal{B}$, ибо функцию, непрерывную на отрезке,

¹⁾ На самом деле вывести их из приведенных выше аксиом нельзя; см. подстрочное примечание на стр. 40. — *Прим. перев.*

можно сколь угодно точно аппроксимировать подходящим образом выбранным полиномом. Плотность одного множества в другом определяется аналогично; каждый элемент второго множества является пределом некоторой последовательности элементов первого множества. Если u — точка пространства $\mathcal{B}$, а $\mathcal{S}$ — некоторое множество точек пространства $\mathcal{B}$, то расстояние $(u, \mathcal{S})$ от u до $\mathcal{S}$ определяется как точная нижняя грань чисел $\|u - v\|$, где v принадлежит $\mathcal{S}$. Можно показать, что если $\mathcal{S}$ замкнуто и u не принадлежит $\mathcal{S}$, то расстояние от u до $\mathcal{S}$ больше нуля (допущение, что оно равно нулю, приводит к противоречию). Если u_0 — элемент пространства $\mathcal{B}$ и r — положительное число, то множество точек u , удовлетворяющих условию $\|u - u_0\| \leq r$, называется *шаром* с центром в точке u_0 радиуса r ; это замкнутое множество.

§ 2.3. Линейные операторы в банаховом пространстве

Действия, выполняемые над функцией при решении краевой задачи, такие, как дифференцирование, интегрирование, составление конечных разностей и умножение на данную функцию, приводят к замене одной функции (или системы функций) другой. В результате происходит преобразование банахова пространства в себя или одной части банахова пространства в другую его часть. Вообще под *преобразованием* T в пространстве $\mathcal{B}$ понимают однозначное сопоставление элементам множества $\mathcal{D} = \mathcal{D}(T)$, называемого *областью определения* T , элементов множества $\mathcal{R} = \mathcal{R}(T)$, называемого *областью значений* преобразования T . Иногда рассматриваются преобразования одного пространства в другое, но нас интересует тот случай, когда $\mathcal{R}$ и $\mathcal{D}$ — множества элементов одного и того же пространства $\mathcal{B}$. Преобразования T и T' называются *идентичными* или *равными* тогда и только тогда, когда совпадают области их определения и когда для каждого u из этой общей области определения $Tu = T'u$.

Оператор, или *операция*, — это то же самое, что преобразование; эти три термина являются равноправными, хотя часто используют термин «преобразование», если речь идет об абстрактном банаховом пространстве, а термин «оператор» — если речь идет о том функциональном пространстве, которое представляет это банахово пространство. Подчеркнем, что при определении оператора необходимо указывать область его определения. Так, нельзя считать оператор d/dx , рассматриваемый на совокупности всех дифференцируемых функций, идентичным с оператором d/dx , область определения которого ограничена, скажем, дважды дифференцируемыми функциями или диффе-

ренцируемыми функциями, обращающимися в нуль при $x = 0$. (Излишне говорить, что символ « d/dx » нельзя использовать, если не указано, какой из этих операторов имеется в виду.) Тем не менее первый из указанных дифференциальных операторов является так называемым расширением двух других. Преобразование T' называется *расширением* преобразования T , если $\mathcal{D}(T)$ содержится в $\mathcal{D}(T')$ и $Tu = T'u$ для всякого u , принадлежащего $\mathcal{D}(T)$. Расширенный оператор производит то же действие, что и исходный, в тех случаях, когда последний может быть применен, но расширенный оператор может быть применен к более широкому классу функций.

Преобразование T называется *линейным* тогда и только тогда, когда при произвольных u и v , принадлежащих $\mathcal{D}(T)$, и произвольных комплексных числах a и b величина $au + bv$ принадлежит $\mathcal{D}(T)$ и $T(au + bv) = aTu + bTv$. Это значит, что не только результат выполнения операции линейно зависит от элемента, над которым выполняется операция, но и область определения преобразования является линейным многообразием.

Линейное преобразование T называется *ограниченным* тогда и только тогда, когда существует такое действительное число K , что $\|Tu\| \leq K\|u\|$ для всех u , принадлежащих $\mathcal{D}(T)$. Наименьшее значение K , для которого это условие выполняется, называется *нормой* преобразования T и обозначается символом $\|T\|$. (Использование такого обозначения основывается на том, что при некоторых условиях множество операторов можно рассматривать как множество элементов некоторого другого банахова пространства, в котором норма элемента определяется как норма соответствующего оператора.) Заметим, что для норм операторов справедливо неравенство треугольника $\|T_1 + T_2\| \leq \|T_1\| + \|T_2\|$.

§ 2.4. Теорема о расширении оператора

Ограниченный линейный оператор T , область определения которого плотна в $\mathcal{B}$, имеет единственное линейное ограниченное расширение T' , область определения которого совпадает с $\mathcal{B}$, такое, что $\|T'\| = \|T\|$.

Действительно, пусть u — произвольный элемент пространства $\mathcal{B}$. По предположению существует последовательность принадлежащих области определения T элементов $u_1, u_2, \dots$, сходящаяся к u . Следовательно,

$$\|Tu_n - Tu_m\| = \|T(u_n - u_m)\| \leq \|T\| \|u_n - u_m\| \rightarrow 0$$

при $n, m \rightarrow \infty$.

Поэтому последовательность элементов $Tu_1, Tu_2, \dots$ фундаментальна и в силу аксиомы о полноте имеет предел. Мы примем

по определению этот предел в качестве значения $T'u$. Читатель без особых затруднений сможет проверить, что (1) этот предел совпадает с Tu , если $u \in \mathcal{D}(T)$, (2) этот предел не зависит от выбора последовательности $u_1, u_2, \dots$ для данного u , (3) этот предел линейно зависит от u и (4) $\|T'\| = \|T\|$. Тем самым теорема доказана.

§ 2.5. Принцип равномерной ограниченности

Класс $\mathcal{C}$ линейных преобразований T , определенных на всем пространстве, называется *равномерно ограниченным*, если существует такое действительное число K , что $\|Tu\| \leq K\|u\|$ для всех $u \in \mathcal{B}$ и всех $T \in \mathcal{C}$. Ясно, что в этом случае все преобразования класса $\mathcal{C}$ являются ограниченными. Но не всякий класс $\mathcal{C}$ ограниченных преобразований равномерно ограничен (если только он не является конечным). Одно условие, обеспечивающее равномерную ограниченность, дается следующей теоремой.

Если $\mathcal{C}$ — некоторый класс ограниченных линейных преобразований, определенных всюду в $\mathcal{B}$, и если для каждого элемента $u \in \mathcal{B}$ существует такое действительное число $K_1(u)$, что $\|Tu\| \leq K_1(u)$ для всех $T \in \mathcal{C}$, то класс $\mathcal{C}$ равномерно ограничен.

Доказательство. Пусть $K_0(T)$ обозначает норму $\|T\|$ оператора T . Нужно доказать, что из неравенств

$$\|Tu\| \leq K_0(T)\|u\|$$

и

$$\|Tu\| \leq K_1(u)$$

следует существование постоянной K_2 , не зависящей ни от u , ни от T , такой, что

$$\|Tu\| \leq K_2\|u\|,$$

причем каждое неравенство удовлетворяется для всех $u \in \mathcal{B}$ и всех $T \in \mathcal{C}$. Рассмотрим последовательность множеств $M_1, M_2, \dots$, где M_p определяется как множество всех таких $u \in \mathcal{B}$, для которых $\|Tu\| \leq p\|u\|$ при всяком $T \in \mathcal{C}$. Наша цель — показать, что одно из этих множеств заполняет все пространство $\mathcal{B}$. Доказательство проведем в несколько шагов.

1. Каждый элемент $u \in \mathcal{B}$ принадлежит хотя бы одному из множеств M_p , так как $u = 0$ принадлежит всем M_p , а $u \neq 0$ принадлежит тем M_p , для которых $p \geq K_1(u)/\|u\|$.

2. Каждое из множеств M_p замкнуто. Действительно, если $u_1, u_2, \dots$ — последовательность элементов из M_p , для которой u является пределом, и если ε — произвольное положительное число, то мы можем выбрать n столь большое, что $\|u - u_n\| < \varepsilon$, так что $\|Tu - Tu_n\| \leq K_0(T)\varepsilon$. Следовательно, в силу неравен-

ства треугольника

$$\begin{aligned}\|Tu\| &\leq \|Tu_n\| + K_0(T)\varepsilon \leq \\ &\leq p\|u_n\| + K_0(T)\varepsilon \leq \\ &\leq p\|u\| + p\varepsilon + K_0(T)\varepsilon.\end{aligned}$$

Так как ε произвольно, то отсюда следует, что $\|Tu\| \leq p\|u\|$ и, значит, $u \in M_p$; тем самым M_p представляет собой замкнутое множество.

3. По меньшей мере одно из множеств M_p содержит шар. Допустим противное. Пусть S_1 — шар, содержащий все u с $\|u\| \leq 1$. По предположению M_1 не содержит S_1 , поэтому в S_1 существует точка, которая не принадлежит M_1 . Эту точку можно считать лежащей внутри S_1 . В самом деле, если множество M_1 содержит все внутренние точки S_1 , то оно содержит и границу S_1 , ибо M_1 — замкнутое множество. Вокруг одной из таких точек построим шар S_2 радиуса не большего $1/2$ и столь малого, что S_2 целиком лежит в S_1 и не содержит точек из M_1 ; это возможно потому, что избранная точка находится на положительном расстоянии как от замкнутого множества M_1 , так и от поверхности шара S_1 , точки которой также образуют замкнутое множество (в соответствии со сказанным в § 2.2). Аналогично построим шар S_3 радиуса не большего $1/4$, лежащий в S_2 и вне M_2 , и т. д. Так построим для каждого n шар S_n радиуса не большего $2^{-(n-1)}$, лежащий в S_{n-1} и вне M_{n-1} . Центры этих шаров образуют фундаментальную последовательность, и легко видеть, что предел этой последовательности принадлежит всем шарам $S_1, S_2, \dots$ и поэтому не может принадлежать ни одному из множеств M_p . Но это противоречит первому пункту доказательства; следовательно, наше допущение ложно.

4. Пусть натуральное число p таково, что M_p содержит шар, и пусть u_0 и r — соответственно центр и радиус этого шара. Если $u \neq 0$ — произвольная точка пространства B , то точка

$$g = u_0 + \frac{r}{\|u\|} u$$

принадлежит поверхности указанного шара и, следовательно, множеству M_p . Поэтому

$$\|Tg\| \leq p\|g\| \leq p(\|u_0\| + r).$$

Но

$$u = \frac{\|u\|}{r}(g - u_0),$$

так что

$$\|Tu\| \leq \frac{\|u\|}{r}[p(\|u_0\| + r) + \|Tu_0\|].$$

Это неравенство (ясно, что оно имеет место и для $u = \emptyset$) совпадает с тем, которое входит в формулировку теоремы, если положить

$$K_2 = \frac{1}{r} [p(\|u_0\| + r) + K_1\|u_0\|];$$

тем самым теорема доказана.

Нужно заметить следующее. В начале доказательства мы говорили о том, что наша цель — показать, что одно из множеств M_p совпадает со всем банаховым пространством. Вместо того чтобы доказать это непосредственно, мы доказали сначала, что одно из множеств M_p содержит шар, хотя быть может и весьма малый. Это было наиболее трудной частью нашего доказательства. После этого уже легко было доказать без дополнительного применения аксиомы о полноте, что каждое из множеств M_p , начиная с достаточно больших p , совпадает со всем банаховым пространством. Это имеет место для всякого M_p с $p \geq K_2$.

§ 2.6. Основная теорема о сходимости

Следующая теорема, принадлежащая Канторовичу [1948], является абстрактным прототипом всех теорем, устанавливающих эквивалентность устойчивости и сходимости.

Пусть $\mathcal{B}'$ и $\mathcal{B}''$ — два банахова пространства с нормами $\|\cdot\|'$ и $\|\cdot\|''$, и пусть T и T_m ($m = 1, 2, 3, \dots$) — линейные операторы, отображающие $\mathcal{B}'$ на $\mathcal{B}''$. Предположим, что (1) для любого $f \in \mathcal{B}''$ уравнение $Tu = f$ имеет единственное решение $u \in \mathcal{B}'$; (2) каждый оператор T_m имеет ограниченный обратный оператор T_m^{-1} и аппроксимирует T в том смысле, что $\|Tu - T_mu\| \rightarrow 0$ при $m \rightarrow \infty$ для любого $u \in \mathcal{B}'$.

Тогда решение u_m уравнения $T_mu_m = f$ сходится к решению u уравнения $Tu = f$ в том и только в том случае, когда аппроксимация устойчива, т. е. обратные операторы T_m^{-1} равномерно ограничены.

Доказательство. Если $\|T_m^{-1}\| \leq M$, где M не зависит от m , то

$$\|u - u_m\|' = \|T_m^{-1}T_m(u - u_m)\|' \leq M\|T_m(u - u_m)\|''.$$

Но

$$\|T_m(u - u_m)\|'' = \|(T_m - T)u + Tu - T_mu_m\|'' = \|(T_m - T)u\|'' \rightarrow 0$$

при $m \rightarrow \infty$ по предположению. С другой стороны, если для любого f $\|u - u_m\|' \rightarrow 0$ при $m \rightarrow \infty$, то последовательность $\|u_m\|' = \|T_m^{-1}f\|'$ ограничена. В силу принципа равномерной

ограниченности отсюда следует равномерная ограниченность норм операторов T_m^{-1} .

При применении этой теоремы к краевым задачам элементами пространств $\mathcal{B}'$ и $\mathcal{B}''$ должны были бы быть функции от пространственных переменных и времени, однако в данном случае оказывается более удобным учитывать их зависимость только от пространственных переменных, используя соответствующую модификацию основной теоремы. Преимуществом такого подхода является также возможность несколько ослабить предположение (1) этой теоремы.

§ 2.7. Замкнутые операторы

Хотя операторы, которые мы будем рассматривать, не являются в общем случае ограниченными (в частности это относится к дифференциальным операторам), они часто обладают свойством замкнутости. Оператор T с областью определения $\mathcal{D}(T)$ называется *замкнутым*, если для любой последовательности $\{u_j\} \in \mathcal{D}(T)$, сходящейся к некоторому u^* и такой, что $\{Tu_j\}$ также сходится к некоторому w^* , $u^* \in \mathcal{D}(T)$ и $Tu^* = w^*$.

Все рассматриваемые в дальнейшем операторы можно считать замкнутыми: это будет доказано в § 3.6. Грубо говоря, каждый оператор можно сделать замкнутым путем выбора подходящей области определения.

В качестве примера рассмотрим оператор $T = d/dx$, действующий на функции $u(x)$ из банахова пространства $\mathcal{B} = L_2[0, 1]$, т. е. пространства функций, интегрируемых с квадратом на отрезке $[0, 1]$. Если в качестве области определения $\mathcal{D}(T)$ взять множество абсолютно непрерывных функций $u(x)$, $0 \leq x \leq 1$ (абсолютно непрерывная функция имеет производную почти всюду и равна интегралу Лебега от своей производной; см. Шилов, 1961, стр. 287), то T будет замкнутым оператором. Это верно потому, что если функции последовательности $\{u_j(x)\}$ абсолютно непрерывны, а $Tu_j(x)$, $\varphi(x)$ и $\psi(x)$ принадлежат $L_2[0, 1]$ и таковы, что

$$\|u_j - \varphi\| \rightarrow 0, \quad \left\| \frac{d}{dx} u_j - \psi \right\| \rightarrow 0,$$

то функция $\varphi(x)$ абсолютно непрерывна и $d\varphi/dx = \psi$ всюду, где производная существует.

Глава 3

ЛИНЕЙНЫЕ РАЗНОСТНЫЕ УРАВНЕНИЯ

В этой главе мы изложим несколько упрощенный вариант разработанной Питером Лаксом общей теории аппроксимации линейных краевых задач с помощью конечноразностных уравнений.

§ 3.1. Корректно поставленные краевые задачи

Задачи, которые мы будем рассматривать, можно представить в абстрактной форме, используя терминологию банаховых пространств. Пусть требуется найти однопараметрическое семейство элементов $u(t)$ из банахова пространства $\mathcal{B}$ (t — действительный параметр), таких, что

$$\frac{d}{dt} u(t) = Au(t), \quad 0 \leq t \leq T, \quad (3.1)$$

и

$$u(0) = u_0, \quad (3.2)$$

где A — линейный оператор, а u_0 — заданный элемент из $\mathcal{B}$, характеризующий начальное состояние физической системы. Определение производной $du(t)/dt$ будет дано ниже.

Граничные условия, если они существуют, предполагаются линейными и однородными; кроме того, предполагается, что область определения оператора A состоит из функций, удовлетворяющих этим граничным условиям.

Применение изложенной схемы не ограничивается уравнениями первого порядка, так как системы более высокого порядка могут быть сведены к системам первого порядка путем соответственного увеличения числа независимых переменных и числа дифференциальных уравнений. Однако способ понижения порядка является существенным: будет показано, что задача может быть корректно поставленной в указанном ниже смысле при одном выборе функционального пространства и не обладать этим свойством при другом его выборе.

Применение этой схемы не ограничивается также системами дифференциальных уравнений. Излагаемая здесь теория применима, например, к некоторым интегро-дифференциальным уравнениям.

У самого Лакса (в докладе на семинаре Нью-Йоркского университета, 1953 г.) допускалось, что оператор A может явно зависеть от параметра t , как в случае, когда скажем коэффициенты дифференциального уравнения зависят от времени. Эта теория почти идентична представленной здесь, но требует несколько более громоздкой записи.

Уточняя постановку задачи (3.1), (3.2), определим *точное решение* этой задачи как такое однопараметрическое семейство $u(t)$, что $u(t)$ принадлежит области определения оператора A при любом t из отрезка $0 \leq t \leq T$, $u(0) = u_0$ и

$$\left\| \frac{u(t + \Delta t) - u(t)}{\Delta t} - Au(t) \right\| \rightarrow 0 \quad \text{при} \quad \Delta t \rightarrow 0, \quad 0 \leq t \leq T. \quad (3.3)$$

Задача (3.1), (3.2) считается корректно поставленной, если совокупность ее точных решений достаточно обширна и если эти решения в некотором смысле однозначно и непрерывно зависят от начальных значений. (Вопрос о том, является ли некоторый заданный оператор A таким, что соответствующая краевая задача корректно поставлена в сформулированном выше смысле, будет рассматриваться не в плане общей теории, а лишь в связи с конкретными примерами.)

Пусть $\mathcal{D}$ — множество таких элементов u_0 из $\mathcal{B}$, для каждого из которых существует единственное точное решение задачи (3.1), (3.2) с $u(0) = u_0$, причем сходимость в (3.3) равномерна по t . (Ниже мы увидим, что эта сходимость автоматически оказывается равномерной для любого точного решения рассматриваемых в этой главе корректно поставленных задач. Однако при проверке корректности постановки конкретной задачи необходимо выяснять вопрос о единственности ее точного решения для достаточно представительного множества начальных значений, решение которого значительно упрощается, если заранее считать эту сходимость равномерной.) Для каждого фиксированного t соответствие между u_0 и $u(t)$ определяет преобразование в $\mathcal{B}$ с областью определения $\mathcal{D}$. Легко проверить, что это преобразование линейно и что, если обозначить его через $E_0(t)$, формула

$$u(t) = E_0(t) u_0 \quad \text{для} \quad u_0 \in \mathcal{D}$$

даст решение задачи (3.1), (3.2) для тех начальных состояний, для которых точное решение существует. В некоторых случаях можно написать явную формулу для $E_0(t)$ (например, интегральное выражение для решения линейной задачи теплопроводности, но это не является необходимым для наших рассуждений).

Задача (3.1), (3.2), определяемая линейным оператором A , называется *корректно поставленной* (это понятие принадлежит

Адамару; впоследствии появилось несколько его вариантов, слегка различающихся между собой), если выполнены следующие два условия:

1. Область определения $\mathcal{D}$ преобразования $E_0(t)$ плотна в $\mathcal{B}$.

2. Семейство преобразований $E_0(t)$ равномерно ограничено, т. е. существует такое K , что $\|E_0(t)\| \leq K$ при $0 \leq t \leq T$.

Второе из этих условий в действительности означает, что решение задачи (3.1), (3.2) непрерывно зависит от начальных значений: в самом деле, если $u(t)$ и $v(t)$ — точные решения, соответствующие начальным элементам u_0 и v_0 , то $\|u(t) - v(t)\| \leq K\|u_0 - v_0\|$, так что если начальные состояния мало отличаются одно от другого, то и последующие состояния, взятые в один и тот же момент времени, близки другу к другу. Первое же условие утверждает, что если для некоторого начального элемента u_0 точное решение не существует, то мы можем аппроксимировать этот начальный элемент сколь угодно точно посредством таких начальных элементов, для которых точные решения существуют. В качестве примера можно снова рассмотреть задачу теплопроводности: если начальная температура распределена разрывно, то мы аппроксимируем ее сходящейся к ней последовательностью непрерывных (и даже дважды дифференцируемых) начальных температур; последовательность построенных таким образом решений сходится к функции, которая может быть интерпретирована как решение задачи теплопроводности с разрывным распределением начальной температуры.

В соответствии с теоремой о расширении из § 2.4 ограниченное линейное преобразование $E_0(t)$ с плотной в $\mathcal{B}$ областью определения имеет расширение $E(t)$, называемое *обобщенным разрешающим оператором*, определенное на всем $\mathcal{B}$ и ограниченное по норме тем же K , что и $E_0(t)$. Равенство

$$u(t) = E(t)u_0$$

дает обобщенное решение задачи (3.1), (3.2) для произвольного начального элемента $u_0 \in \mathcal{B}$.

Семейство разрешающих операторов является *полугруппой*:

$$E(s+t) = E(s)E(t)$$

при $s \geq 0$, $t \geq 0$. Хотя этот факт уже неверен в более общем случае, когда A зависит от t , и несуществен для излагаемой теории, мы будем использовать его всякий раз, когда это может упростить доказательство.

Для того чтобы краевая задача для уравнения в частных производных в ограниченной области была корректно поставлена, нужно, чтобы граничные условия были подходящими, ибо

от них зависит область определения оператора A и, следовательно, операторов $E_0(t)$. Рассмотрим для примера простейшую задачу теплопроводности на отрезке $a \leq x \leq b$. Если граничных условий слишком много, скажем $u(x, t) = 0$ и $(\partial/\partial x)u(x, t) = 0$ при $x = a$ и $x = b$, то вообще не существует точных решений, кроме тождественного нуля, и потому $\mathcal{D}$ не плотно в $\mathcal{B}$. С другой стороны, если граничных условий слишком мало, например их нет совсем, то ни для одного начального значения нет *единственности* точного решения, и $\mathcal{D}$ снова не плотно в $\mathcal{B}$.

Из (3.3) следует, что точное решение является непрерывной функцией от t в смысле нормы пространства $\mathcal{B}$: $\|u(t + \Delta t) - u(t)\| \rightarrow 0$ при $\Delta t \rightarrow 0$, $0 \leq t \leq T$; применяя неравенство треугольника, можно установить непрерывность и обобщенного решения для $0 \leq t \leq T$.

Если $u_0 \in \mathcal{D}$, то $(1/\delta)[E(\delta) - I]u_0$ аппроксимирует Au_0 ; учитывая перестановочность $E(t)$ с $E(\delta)$, получаем отсюда после перехода к пределу, что

$$E(t)Au_0 = AE(t)u_0.$$

В силу этого равенства левую часть соотношения (3.3) можно переписать в виде

$$\left\| E(t) \left\{ \frac{E(\Delta t) - I}{\Delta t} - A \right\} u_0 \right\|;$$

поскольку $\|E(t)\| \leq K$ при $0 \leq t \leq T$, отсюда следует, что для всякого точного решения корректно поставленной задачи сходимость в (3.3) равномерна по t .

В случае когда оператор A зависит явным образом от времени, обобщенный разрешающий оператор становится функцией двух переменных; действительно, если в момент времени t_0 задано начальное состояние u_0 , то $u(t) = E(t, t_0)u_0$, причем в силу полугруппового свойства $E(t_2, t_0) = E(t_2, t_1)E(t_1, t_0)$ при $t_0 \leq t_1 \leq t_2$.

§ 3.2. Конечноразностные аппроксимации

Согласно замечанию, сделанному в конце § 2.1, аппроксимации, получаемые посредством конечноразностных уравнений, также можно представлять точками пространства $\mathcal{B}$. При этом однопараметрическое семейство $u(t)$ заменяется последовательностью точек $u^0, u^1, u^2, \dots$, причем u^n по предположению есть приближенное значение $u(n\Delta t)$, а Δt обозначает малое приращение.

Соответствующие конечноразностные уравнения можно записать в виде

$$B_1 u^{n+1} = B_0 u^n, \quad (3.4)$$

где $B_0 = B_0(\Delta t, \Delta x_1, \Delta x_2, \dots)$ и $B_1 = B_1(\Delta t, \Delta x_1, \Delta x_2, \dots)$ — линейные конечноразностные операторы, зависящие, как это видно из обозначений, от величин конечных приращений $\Delta t, \Delta x_1, \Delta x_2, \dots$, а также возможно от пространственных переменных. Предположим, что B_0 и B_1 , как и оператор A , не зависят от t . В каждой точке пространства обе части уравнения (3.4) являются линейными комбинациями значений функции u на некотором множестве близких друг другу точек. Мы предположим, что всегда имеется некоторый способ для вычисления u^{n+1} при любом u^n , причем u^{n+1} однозначно и непрерывно зависит от u^n . Таким образом, мы предполагаем, что оператор B_1^{-1} существует, оператор $B_1^{-1}B_0$ ограничен и оба они определены на всем пространстве $\mathcal{B}$. Эти предположения будут очевидным образом выполнены для любой разумной разностной схемы.

Формулы вида (3.4) называются *двухслойными*, так как они содержат всего два «временных слоя» t^n и t^{n+1} . Нужно заметить, однако, что применение формул такого рода не означает сужения класса решаемых задач до систем дифференциальных уравнений первого порядка по времени. Как отмечалось выше, системы дифференциальных уравнений высшего порядка могут быть сведены к системам первого порядка путем введения новых зависимых переменных. Введение этих новых переменных сведет соответствующую конечноразностную систему к двухслойной, если число входящих в нее слоев не превышает порядка по t первоначальной системы дифференциальных уравнений. (Некоторые трехслойные формулы для уравнения диффузии обсуждаются в гл. 8; см. схемы 7—11 и 13 из табл. 8.1.) Многослойные формулы будут рассмотрены в гл. 7. Иногда используются термины «одношаговый» и «многошаговый».

Понятия устойчивости и сходимости, с которыми мы имеем дело, предполагают бесконечную последовательность вычислений при безграничном измельчении ячеек сетки. Пусть выполнены соотношения

$$\Delta x_i = g_i(\Delta t), \quad i = 1, 2, \dots, d,$$

где d — размерность пространства, показывающие, как стремятся к нулю пространственные приращения при приближении к нулю временного приращения. Положим

$$B_1^{-1}(\Delta t, g_1(\Delta t), g_2(\Delta t), \dots) B_0(\Delta t, g_1(\Delta t), g_2(\Delta t), \dots) = C(\Delta t),$$

так что

$$u^{n+1} = C(\Delta t) u^n. \quad (3.5)$$

Мы приходим теперь к условию согласованности. Так как отношение

$$\frac{u^{n+1} - u^n}{\Delta t}$$

является приближенным значением производной по времени, то отношение

$$\frac{C(\Delta t)u - u}{\Delta t}$$

должно в некотором смысле аппроксимировать величину Au . Мы не можем ожидать того, что это условие имеет место для всех $u \in \mathcal{B}$, хотя бы потому, что в общем случае оператор A определен не на всем $\mathcal{B}$. Но мы хотим, чтобы оно выполнялось почти для всех u , соответствующих точным решениям нашей задачи. Будем говорить, что семейство операторов $C(\Delta t)$ *согласованно аппроксимирует* краевую задачу (3.1), (3.2), если для любого $u(t)$ из некоторого класса $\mathcal{U}$ точных решений, начальные элементы которых образуют в $\mathcal{B}$ плотное множество, выполняется предельное соотношение

$$\left\| \left\{ \frac{C(\Delta t) - I}{\Delta t} - A \right\} u(t) \right\| \rightarrow 0 \quad \text{при } \Delta t \rightarrow 0, \quad 0 \leq t \leq T, \quad (3.6a)$$

называемое *условием согласованности*. Здесь I обозначает единичный оператор.

Поскольку это условие относится к точным решениям, то, комбинируя его с (3.3), получаем эквивалентное (3.6a) соотношение

$$\left\| \frac{u(t + \Delta t) - C(\Delta t)u(t)}{\Delta t} \right\| \rightarrow 0 \quad \text{при } \Delta t \rightarrow 0, \quad 0 \leq t \leq T. \quad (3.6b)$$

Стоящая здесь под знаком нормы величина, называемая *погрешностью аппроксимации*, показывает, насколько хорошо точное решение краевой задачи удовлетворяет конечноразностным уравнениям. Предположим, что сходимость в (3.6a) и (3.6b) равномерна по t ; тогда для любого $\varepsilon > 0$ существует такое $\delta > 0$, что

$$\| \{C(\Delta t) - E(\Delta t)\} u(t) \| < \varepsilon \Delta t \quad \text{при } 0 \leq t \leq T, \quad 0 < \Delta t < \delta. \quad (3.6c)$$

§ 3.3. Сходимость

Применив n раз оператор $C(\Delta t)$ к u_0 , получим величину $u^n = C(\Delta t)^n u_0$, которая по предположению аппроксимирует $u(n\Delta t)$. Так как $u(t) = E(t)u_0$, то мы введем следующее определение. Семейство операторов $C(\Delta t)$ обеспечивает *сходящуюся аппроксимацию* задачи (3.1), (3.2), если при любом фиксированном t из отрезка $0 \leq t \leq T$ для каждого $u_0 \in \mathcal{B}$ и для каждой

сходящейся к нулю последовательности положительных приращений времени $\Delta_1 t, \Delta_2 t, \dots, \Delta_j t, \dots$ имеет место предельное соотношение

$$\|C(\Delta_j t)^{n_j} u_0 - E(t) u_0\| \rightarrow 0 \quad \text{при } j \rightarrow \infty, \quad (3.7)$$

где n_j — целое число, близкое к $t/\Delta_j t$ в том смысле, что $n_j \Delta_j t \rightarrow t$ при $j \rightarrow \infty$.

§ 3.4. Устойчивость

Если в последовательности вычислений при $\Delta_j t \rightarrow 0$ каждое вычисление проводится от $t = 0$ до $t = T$, то каждый из бесконечного множества операторов

$$C(\Delta_j t)^n, \quad \begin{matrix} 0 \leq n \Delta_j t \leq T, \\ j = 1, 2, 3, \dots, \end{matrix} \quad (3.8)$$

применяется к u_0 . Суть устойчивости состоит в том, что существуют пределы, которых не могут превзойти компоненты начальной функции, преобразующиеся в процессе вычислений. Поэтому аппроксимацию $C(\Delta_j t)$ называют *устойчивой*, если при некотором $\tau > 0$ бесконечное множество операторов

$$C(\Delta t)^n, \quad \begin{matrix} 0 < \Delta t < \tau, \\ 0 \leq n \Delta t \leq T, \end{matrix} \quad (3.9)$$

равномерно ограничено.

Будем считать в дальнейшем, что конечноразностный оператор непрерывно зависит от Δt , если Δt положительно, и не превосходит некоторого $\tau > 0$. В силу этого предположения коэффициенты разностного оператора, так же как и его норма, будут непрерывными функциями от Δt . В этом случае множество операторов (3.9) автоматически будет равномерно ограниченным на любом отрезке $\tau' \leq \Delta t \leq \tau$, где $\tau' > 0$; существо данного здесь определения устойчивости состоит в том, что это множество операторов остается равномерно ограниченным, когда $\tau' \rightarrow 0$.

Понятие устойчивости не имеет никакого отношения к дифференциальным уравнениям, подлежащим численному интегрированию, и является исключительно свойством системы разностных уравнений.

§ 3.5. Теорема Лакса об эквивалентности

Пусть дана корректно поставленная задача (3.1), (3.2), и пусть ее конечноразностная аппроксимация удовлетворяет условию согласованности; тогда устойчивость необходима и достаточна для сходимости.

В соответствии с определением, данным в § 3.3, здесь имеется в виду сходимость при произвольном начальном элементе u_0 . Как замечено в первой главе (см. рассуждения по поводу неравенства (1.15)), и неустойчивая схема может иногда обеспечивать сходимость при специально подобранных начальных элементах.

Сначала мы покажем, что сходящаяся схема необходимо является устойчивой. Мы утверждаем, что для всякой сходящейся схемы и для произвольного начального фиксированного элемента $u_0 \in \mathcal{B}$ величины

$$\|C(\Delta t)^n u_0\|, \quad \begin{array}{l} 0 < \Delta t < \tau, \\ 0 \leq n \Delta t \leq T, \end{array}$$

ограничены при некотором $\tau > 0$. Действительно, если это не так, то найдутся две последовательности $\Delta_1 t, \Delta_2 t, \dots, \Delta_j t, \dots$ и $n_1, n_2, \dots, n_j, \dots$, для которых нормы элементов $C(\Delta_1 t)^{n_1} u_0, C(\Delta_2 t)^{n_2} u_0, \dots, C(\Delta_j t)^{n_j} u_0, \dots$, неограниченно возрастают (при этом $\Delta_j t$ должны стремиться к нулю в силу предположения о непрерывной зависимости $C(\Delta t)$ от положительных значений Δt); из этих элементов мы можем выбрать подпоследовательность, для которой величины $n_j \Delta_j t$ сходятся к некоторому t из отрезка $0 \leq t \leq T$; но это противоречит предположению о сходимости схемы, поскольку при наличии сходимости нормы элементов этой подпоследовательности должны были бы стремиться к конечному пределу $\|E(t)u_0\|$. Следовательно, существует такая функция $K_1(u)$, что неравенство $\|C(\Delta t)^n u\| \leq K_1(u)$ выполняется для всех операторов из множества (3.9) и всех $u \in \mathcal{B}$. Но тогда в силу принципа равномерной ограниченности множество (3.9) равномерно ограничено. Таким образом, аппроксимация устойчива в смысле определения § 3.4.

Чтобы доказать обратное утверждение, предположим, что $u(t) = E(t)u_0$ является точным решением, принадлежащим классу $\mathcal{U}$, о котором шла речь при определении согласованности. Пусть ε и δ те же, что и в условии согласованности в форме (3.6с), n_j и $\Delta_j t$ выбраны так же, как и при определении сходимости, а ψ_j обозначает разность между вычисленным и точным значением u в момент времени $n_j \Delta_j t$, т. е.

$$\begin{aligned} \psi_j &= [C(\Delta_j t)^{n_j} - E(n_j \Delta_j t)] u_0 = \\ &= \sum_{k=0}^{n_j-1} C(\Delta_j t)^k [C(\Delta_j t) - E(\Delta_j t)] E((n_j - 1 - k) \Delta_j t) u_0. \end{aligned} \quad (3.10)$$

Третья часть этого равенства, как легко видеть, совпадает со второй: после приведения подобных остаются только первый

и последний члены. Норму величины ψ_j можно оценить с помощью (3.6с) и неравенства треугольника: если $0 < \Delta_j t < \delta$, то

$$\|\psi_j\| \leq K_2 \sum_{k=0}^{n_j-1} \varepsilon \Delta_j t = K_2 \varepsilon n_j \Delta_j t \leq K_2 \varepsilon T, \quad (3.11)$$

где K_2 обозначает равномерную границу множества (3.9). Так как ε произвольно, то $\|\psi_j\| \rightarrow 0$ при $\Delta_j t \rightarrow 0$, $n_j \Delta_j t \rightarrow t$. Для доказательства сходимости покажем, что в пределе при $j \rightarrow \infty$ в (3.10) можно заменить $E(n_j \Delta_j t)$ на $E(t)$. Если $s = |t - n_j \Delta_j t|$, $t' = \min\{t, n_j \Delta_j t\}$, то в силу полугруппового свойства семейства $E(t)$ имеем $E(s + t') = E(s)E(t')$, так что $E(n_j \Delta_j t) - E(t) = \pm [E(s) - I]E(t')$, причем знак определяется знаком разности $t - n_j \Delta_j t$. В любом случае

$$\| [E(n_j \Delta_j t) - E(t)] u_0 \| \leq K_E \| [E(s) - I] u_0 \|,$$

где K_E обозначает границу для $\|E(t)\|$ при $0 \leq t \leq T$. Но правая часть последнего неравенства стремится к нулю, если $s \rightarrow 0$, т. е. если $j \rightarrow \infty$. Следовательно, величина $\| [C(\Delta_j t)^{n_j} - E(t)] u_0$ может быть сделана сколь угодно малой выбором достаточно малых $\Delta_j t$ и $|t - n_j \Delta_j t|$. Это справедливо для любого u_0 , являющегося начальным элементом точного решения из класса $\mathcal{U}$: но такие элементы плотны в $\mathcal{B}$, так что для любого $u \in \mathcal{B}$ из них можно выбрать последовательность $u_1, u_2, \dots$, сходящуюся к u . Поэтому

$$\begin{aligned} [C(\Delta_j t)^{n_j} - E(t)]u &= [C(\Delta_j t)^{n_j} - E(t)]u_m + \\ &+ C(\Delta_j t)^{n_j}(u - u_m) - E(t)(u - u_m). \end{aligned} \quad (3.12)$$

Здесь два последних члена в правой части могут быть сделаны сколь угодно малыми посредством выбора достаточно большого m , поскольку класс (3.9) и множество операторов $E(t)$ равномерно ограничены, а малость первого члена может быть обеспечена, как уже отмечалось выше, за счет выбора достаточно малых $\Delta_j t$ и $|t - n_j \Delta_j t|$. Так как u — произвольный элемент из $\mathcal{B}$, то тем самым сходимость установлена и теорема об эквивалентности доказана.

Отметим в качестве естественного следствия этой теоремы, что для данного начального элемента u_0 сходимость равномерна по t на отрезке $0 \leq t \leq T$ в том смысле, что ограничения, которые нужно налагать на $\Delta_j t$ и на $|t - n_j \Delta_j t|$, чтобы сделать (3.12) произвольно малым, не зависят ни от выбора t , ни от выбора последовательности $\Delta_j t$. Это обстоятельство имеет большое практическое значение, так как позволяет при численном интегриро-

вании находить такой шаг Δt , при котором приближенное решение оказывается достаточно точным на всем отрезке $0 \leq t \leq T$ сразу. Часто для достижения нужной точности Δt варьируют в процессе вычислений, но существует такое предельное положительное значение Δt , ниже которого заходить нет смысла.

Изложенную теорему Лакса можно очень просто применить к неявному разностному уравнению (1.21) для одномерной задачи диффузии

$$\frac{u_j^n - u_j^{n-1}}{\Delta t} = \sigma \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2}.$$

Прежде всего покажем, что решение этого уравнения удовлетворяет следующему принципу максимума¹⁾. Предположим, что уравнение рассматривается в прямоугольнике $0 \leq x \leq a$, $0 \leq t \leq T$ и что Δx выбирается равным a/J , где J — натуральное число. Тогда максимальное значение M_1 , достигаемое величиной u_j^n внутри этого прямоугольника, не может превышать максимального значения M_2 , достигаемого начальными и граничными значениями (т. е. значениями на отрезках прямых $t = 0$, $x = 0$, $x = a$). Действительно, допустим противное, а именно что $M_1 > M_2$, и пусть (n, j) — первая внутренняя точка сетки, в которой $u_j^n = M_1$ (первая в том смысле, что индексы n и j имеют наименьшие значения). Тогда непосредственно ясно, что в этой точке сетки левая часть приведенного выше уравнения должна быть положительна, а правая отрицательна, ибо u_j^n по предположению превосходит соседнее значение u_{j-1}^n слева и соседнее значение u_j^{n-1} снизу и по меньшей мере равно соседнему значению u_{j+1}^n справа. Следовательно, наше предположение ложно, и принцип максимума установлен. Очевидно, что эти соображения можно применить и к $-u_j^n$ и тем самым установить, что $|u_j^n|$ ограничены при любом выборе сетки, т. е., другими словами, что рассматриваемое разностное уравнение устойчиво.

Эти соображения могут быть применены также при доказательстве устойчивости определенного аналогичным образом разностного уравнения для более общей задачи

$$\frac{\partial u}{\partial t} = a(x) \frac{\partial}{\partial x} \left[b(x) \frac{\partial u}{\partial x} \right], \quad a(x) > 0, \quad b(x) > 0$$

(проведение доказательства предоставляется читателю). Условие согласованности в этом случае также выполнено и сходимость тем самым обеспечена независимо от того, каким образом $\Delta t \rightarrow 0$ и $\Delta x \rightarrow 0$.

¹⁾ Дж. Франклин, частное сообщение [1954]; Дж. Дуглас мл. [1955] П. Лаасонен [1949].

§ 3.6. Замкнутый оператор A'

В этом параграфе будет показано, что если оператор A в (3.1) определяет корректно поставленную краевую задачу, то существует замкнутый оператор A' , который также определяет корректно поставленную краевую задачу, причем каждое точное решение первой задачи является также точным решением второй задачи и обе задачи имеют один и тот же обобщенный разрешающий оператор $E(t)$, так что с практической точки зрения они эквивалентны. Оператор A' будет играть важную роль в следующем параграфе, посвященном неоднородным задачам.

Грубо говоря, A' определяется как производная по t от $E(t)$ при $t = 0$. Будем считать, что $\mathcal{D}(A')$ состоит из таких элементов $u \in \mathcal{B}$, для которых в $\mathcal{B}$ существует предел

$$w = \lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} [E(\Delta t) - I] u;$$

для таких u положим $A'u = w$. Ясно, что $\mathcal{D}(A')$ есть линейное многообразие в $\mathcal{B}$, а A' — линейный оператор. Сравнивая приведенное в этом определении равенство с (3.3), находим, что $A'u = Au$ для любого u , соответствующего точному решению исходной задачи.

В соответствии с терминологией теории линейных операторов (см. Данфорд и Шварц, т. 1, 1962), A' называется *инфинитезимальным оператором* семейства операторов $E(t)$, которое является сильно непрерывной полугруппой.

Так как $E(t)$ — ограниченный оператор, перестановочный с $E(\Delta t)$, то

$$E(t) (1/\Delta t) [E(\Delta t) - I] u_0 = (1/\Delta t) [E(\Delta t) - I] E(t) u_0;$$

при $u_0 \in \mathcal{D}(A')$ существует предел левой части при $\Delta t \rightarrow 0$, откуда в силу определения оператора A' следует, что $E(t)u_0 \in \mathcal{D}(A')$ и $E(t)A'u_0 = A'E(t)u_0$, т. е. A' и $E(t)$ перестановочны на элементе u_0 . Таким образом, функция $u(t) = E(t)u_0$ удовлетворяет предельному соотношению

$$\left\| \frac{u(t + \Delta t) - u(t)}{\Delta t} - A'u(t) \right\| \rightarrow 0 \quad \text{при} \quad \Delta t \rightarrow 0,$$

т. е. является решением краевой задачи

$$\frac{d}{dt} u(t) = A'u(t), \quad u(0) = u_0. \quad (3.13)$$

Покажем, что и обратно, единственным решением задачи (3.13) является функция $u(t) = E(t)u_0$. Если $u(t)$ — решение

этой задачи, то при фиксированном $t > 0$ функция

$$g(s) = E(t - s)u(s), \quad 0 \leq s \leq t,$$

постоянна. Действительно,

$$\left. \frac{d}{ds} g(s) \right|_{s=s_0} = \left. \frac{d}{ds} E(t - s)u(s_0) \right|_{s=s_0} + \left. \frac{d}{ds} E(t - s_0)u(s) \right|_{s=s_0};$$

первое слагаемое в правой части можно переписать в виде

$$-\left. \frac{d}{dw} E(w)u(s_0) \right|_{w=t-s_0} = -A'E(t - s_0)u(s_0),$$

потому что функция $E(t)u(s_0)$ удовлетворяет дифференциальному уравнению (3.13); для второго слагаемого имеем

$$E(t - s_0) \left. \frac{d}{ds} u(s) \right|_{s=s_0} = E(t - s_0) A'u(s_0),$$

поскольку $E(t - s_0)$ является ограниченным оператором. Так как A' и $E(t - s_0)$ перестановочны, то $dg(s)/ds = 0$ и потому $g(s)$ является константой. Следовательно, $g(t) = g(0)$, или $u(t) = E(t)u_0$, что и требовалось доказать. Отсюда следует, что (3.13) — корректно поставленная задача.

Покажем теперь, что A' — замкнутый оператор в смысле определения, данного в § 2.7. Сначала напомним некоторые сведения об интегрировании в банаховом пространстве. Пусть совокупность элементов $w(s)$ из $\mathcal{B}$ непрерывно зависит от параметра s . Тогда интеграл Римана

$$I = \int_{s_1}^{s_2} w(s) ds$$

определяется как предел интегральных сумм

$$\sum_j w(s'_j)(s_{j+1} - s_j), \quad s_j \leq s'_j \leq s_{j+1},$$

аналогично тому, как это делается для непрерывной функции действительного переменного. Предлагаем читателю в качестве упражнения проверить с помощью неравенства треугольника, что этот предел единствен и что определенный таким образом интеграл обладает всеми обычными свойствами, например что

$$\frac{dI}{ds_2} = w(s_2),$$

$$\|I\| \leq \int_{s_1}^{s_2} \|w(s)\| ds \leq |s_2 - s_1| \max_{s_1 \leq s \leq s_2} \|w(s)\|.$$

Проверим, что если $u_0 \in \mathcal{D}(A')$, то

$$u(t) - u_0 = \int_0^t A' u(s) ds, \quad (3.14)$$

где $u(t) = E(t)u_0$; это согласуется с уравнением $du/dt = A'u$. Уравнение (3.13) показывает, что для данного $\varepsilon > 0$ существует такое $\delta > 0$, что

$$\left\| \frac{E(\delta) - I}{\delta} u_0 - A'u_0 \right\| < \varepsilon,$$

или, после применения оператора $E(j\delta)$,

$$\| E((j+1)\delta) u_0 - E(j\delta) u_0 - \delta E(j\delta) A'u_0 \| < \varepsilon \delta K,$$

где $K = \max \|E(t)\|$. Полагая $n\delta = t$, суммируя по j выражения, стоящие под знаком нормы в левой части последнего неравенства и оценивая затем эту сумму по норме с помощью неравенства треугольника, получаем

$$\left\| E(t) u_0 - u_0 - \sum_{j=0}^{n-1} \delta E(j\delta) A'u_0 \right\| < \varepsilon t K.$$

Так как функция $E(s)A'u_0$ непрерывна, то сумма в левой части этого неравенства аппроксимирует интеграл из правой части равенства (3.14). В силу произвольности ε отсюда следует справедливость равенств (3.14).

Предположим теперь, что последовательность $u_1, u_2, \dots, u_j, \dots$ элементов из $\mathcal{D}(A')$ сходится к некоторому $u^* \in \mathcal{B}$, а $\{A'u_j\}$ сходится к некоторому w^* из $\mathcal{B}$, и покажем, что $u^* \in \mathcal{D}(A')$ и $w^* = A'u^*$, т. е. что оператор A' замкнут. В самом деле,

$$\begin{aligned} E(\delta) u^* - u^* &= \lim_{j \rightarrow \infty} [E(\delta) u_j - u_j] = \lim_{j \rightarrow \infty} \int_0^\delta A' E(s) u_j ds = \\ &= \lim_{j \rightarrow \infty} \int_0^\delta E(s) A' u_j ds = \int_0^\delta E(s) \lim_{j \rightarrow \infty} A' u_j ds \end{aligned}$$

(предельный переход перестановочен с интегрированием, потому что сходимость здесь равномерна в силу равномерной ограниченности $E(s)$ по s). Следовательно,

$$\frac{E(\delta) u^* - u^*}{\delta} = \frac{1}{\delta} \int_0^\delta E(s) w^* ds,$$

причем правая часть этого равенства стремится к $E(0)w^* = w^*$ при $\delta \rightarrow 0$, так как операция интегрирования непрерывна. Таким образом, замкнутость оператора A' доказана.

Методом, предложенным Гельфандом и Гордингом (Лионс, частное сообщение), можно показать, что области определения оператора A' и всех его степеней плотны в $\mathcal{B}$. Пусть $\varphi(t)$ — бесконечно дифференцируемая функция, равная нулю вне интервала $0 < t < t_0$ (в частности, эта функция обращается в нуль при $t = 0$ вместе со всеми своими производными). Для произвольного элемента $u \in \mathcal{B}$ положим

$$T(\varphi)u = \int_0^\infty E(t)u\varphi(t)dt.$$

Это равенство определяет для каждой допустимой функции φ оператор $T(\varphi)$, заданный на всем пространстве $\mathcal{B}$. Покажем, что $T(\varphi)u$ всегда принадлежат множеству $\mathcal{D}(A')$. Действительно,

$$E(s)T(\varphi)u = \int_0^\infty E(s+t)u\varphi(t)dt = \int_0^\infty E(t)u\varphi(t-s)dt,$$

и потому при $s \rightarrow 0$

$$\frac{E(s)T(\varphi)u - T(\varphi)u}{s} = \int_0^\infty E(t)u \frac{\varphi(t-s) - \varphi(t)}{s} dt \rightarrow T(-\varphi')u,$$

где $\varphi' = d\varphi/dt$. Следовательно, в силу определения A' , $T(\varphi)u \in \mathcal{D}(A')$ и $A'T(\varphi)u = T(-\varphi')u$. Повторяя рассуждения $q-1$ раз, находим, что $(A')^q T(\varphi)u$ также существует и равняется $T((-1)^q \varphi^{(q)})u$. Обозначим через $\mathcal{B}_0$ множество элементов вида $T(\varphi)u$ для всех допустимых функций φ и всех $u \in \mathcal{B}$. Если $\{\varphi_m\}$ является δ -образной последовательностью¹⁾ допустимых функций, т. е. сходится в смысле обобщенных функций к дельта-функции Дирака $\delta(t)$, то для произвольного $u \in \mathcal{B}$ в силу непрерывности $E(t)u$ будем иметь $T(\varphi_m)u \rightarrow u$ при $m \rightarrow \infty$. Следовательно, $\mathcal{B}_0$ плотно в $\mathcal{B}$. Но так как $(A')^q T(\varphi)u \in \mathcal{D}((A')^q)$, то $\mathcal{B}_0 \subset \mathcal{D}((A')^q)$ при любом натуральном q ; поэтому множества $\mathcal{D}((A')^q)$ также плотны в $\mathcal{B}$.

§ 3.7. Неоднородные задачи

Основной целью этого параграфа является решение следующей краевой задачи:

$$\begin{aligned} \frac{d}{dt} u(t) - Au(t) &= g(t), \\ u(0) &= u_0, \end{aligned} \tag{3.15}$$

¹⁾ См., например, Гельфанд и Шиллов* [1958], гл. 1. — Прим. перев.

где u_0 и $g(t)$ заданы, причем функция $g(t)$ равномерно непрерывна по t на отрезке $0 \leq t \leq T$ в смысле нормы пространства $\mathcal{B}$. Достаточно было бы и кусочной равномерной непрерывности $g(t)$; провести доказательство для этого случая мы предоставляем читателю. Мы предполагаем, что определяющий корректную однородную краевую задачу оператор A замкнут и что области определения всех его степеней плотны в $\mathcal{B}$ (как было показано в предыдущем параграфе, это предположение не является ограничением).

Мы покажем, что решение задачи (3.15) имеет вид

$$u(t) = E(t)u_0 + \int_0^t E(t-s)g(s)ds \quad (3.16)$$

при следующих дополнительных ограничениях: (1) u_0 и $g(t)$ принадлежат $\mathcal{D}(A)$; (2) $g(t) \in \mathcal{D}(A^2)$; (3) функции $Ag(t)$ и $A^2g(t)$ непрерывны. Функция (3.16) называется *точным решением* задачи (3.15). Если на u_0 и $g(t)$ не наложено никаких ограничений, кроме непрерывности $g(t)$ или даже только существования интеграла в (3.16), то функция (3.16) называется *обобщенным решением* задачи (3.15). Относящиеся сюда результаты были получены Р. Дж. Томпсоном [1964]; он рассмотрел также квазилинейный случай, для которого $g = g(t, u(t))$.

Из ограниченности $E(t)$ и непрерывности $g(t)$ и $E(t)u$ при любом u следует в силу неравенства треугольника, что функция $E(t-s)g(s)$ непрерывна по s и, значит, интегрируема. Наше утверждение относительно функции (3.16) будет доказано, если мы покажем, что

$$\frac{d}{dt} \int_0^t E(t-s)g(s)ds = E(0)g(t) + \int_0^t \frac{d}{dt} E(t-s)g(s)ds. \quad (3.17)$$

Интеграл в правой части этого равенства существует, потому что подынтегральное выражение равно $AE(t-s)g(s) = E(t-s)Ag(s)$ и непрерывно в силу непрерывности $Ag(s)$ и равномерной ограниченности $E(t-s)$. Для внесения производной под знак интеграла, как и для функций действительного аргумента, достаточно, чтобы соответствующее разностное отношение равномерно по s сходилось к производной, т. е. чтобы

$$\left\| \frac{E(\delta) - I}{\delta} E(t-s)g(s) - AE(t-s)g(s) \right\| \rightarrow 0 \quad (3.18)$$

равномерно по s при $\delta \rightarrow 0$. В силу равномерной ограниченности $E(t-s)$ и перестановочности $E(t-s)$ с A это равносильно тому, что

$$\left\| \frac{E(\delta) - I}{\delta} g(s) - Ag(s) \right\| \rightarrow 0$$

равномерно по s при $\delta \rightarrow 0$. При фиксированном s функция $h(t) = E(t)g(s)$ удовлетворяет не только уравнению $dh/dt = Ah$, но по предположению также и уравнению $d^2h/dt^2 = A^2h$. Поэтому, используя дважды (3.14), получаем

$$\begin{aligned} \frac{E(\delta) - I}{\delta} g(s) - Ag(s) &= \frac{1}{\delta} \left[\int_0^\delta Ah(t) dt - \delta Ah(0) \right] = \\ &= \frac{1}{\delta} \int_0^\delta dt \int_0^t A^2 h(w) dw = \frac{1}{\delta} \int_0^\delta dt \int_0^t E(w) A^2 g(s) dw, \end{aligned}$$

так что

$$\left\| \frac{E(\delta) - I}{\delta} g(s) - Ag(s) \right\| \leq \frac{1}{2} \delta \max_{0 \leq w \leq T} \|E(w) A^2 g(s)\|.$$

Но максимум здесь ограничен в силу непрерывности $A^2 g(s)$ и равномерной ограниченности $E(w)$. Этим доказана равномерность сходимости в (3.18) и тем самым справедливость равенства (3.17).

Равенство (3.17) принимает теперь вид

$$\frac{d}{dt} \int_0^t E(t-s) g(s) ds = g(t) + \int_0^t AE(t-s) g(s) ds.$$

Так как в любой интегральной сумме, аппроксимирующей интеграл в правой части этого равенства, оператор A можно вынести за знак суммы и интеграл

$$\int_0^t E(t-s) g(s) ds$$

существует, то в силу замкнутости оператора A его можно вынести также и за знак интеграла. Следовательно,

$$\frac{d}{dt} \int_0^t E(t-s) g(s) ds = g(t) + A \int_0^t E(t-s) g(s) ds,$$

откуда видно, что функция (3.16) является решением краевой задачи (3.15).

Единственность этого решения следует из единственности решения соответствующей однородной задачи, которая по предположению является корректно поставленной.

При рассмотрении обобщенных решений без ограничения общности можно считать, что $u_0 = 0$, поскольку первое слагаемое в (3.16) есть просто обобщенное решение соответствующей однородной задачи. Обозначим через $\mathscr{B}'$ вспомогательное банахово пространство, элементами которого являются всевоз-

можные непрерывные кривые $v(t) \in \mathcal{B}$, определенные при $0 \leq t \leq T$, с нормой

$$\|v(\cdot)\|_{\mathcal{B}'} = \max_{0 \leq t \leq T} \|v(t)\|_{\mathcal{B}}.$$

Отображение

$$F: g(t) \rightarrow u(t) = \int_0^t E(t-s) g(s) ds$$

есть ограниченное линейное преобразование, определенное на всем $\mathcal{B}'$; обозначим через F_0 сужение F на множество тех элементов $g(\cdot)$ из $\mathcal{B}'$, для которых существует определенное выше точное решение задачи (3.15), и покажем, что множество $\mathcal{D}(F_0)$ плотно в $\mathcal{B}'$. Пусть $g(\cdot)$ — произвольный элемент пространства $\mathcal{B}'$. Разобьем отрезок $[0, T]$ на N отрезков длины δ и в каждой точке разбиения аппроксимируем $g(n\delta)$ элементом $\bar{g}^n \in \mathcal{D}(A^2)$. После этого определим $\bar{g}(t)$ как кусочно линейную функцию, построенную по $\bar{g}^n$. Очевидно, $\bar{g}(t) \in \mathcal{D}(A^2)$ при $0 \leq t \leq T$ и функции $\bar{g}(t)$, $A\bar{g}(t)$ и $A^2\bar{g}(t)$ непрерывны. Далее $\|g(t) - \bar{g}(t)\|$ можно сделать равномерно малым на всем отрезке $0 \leq t \leq T$ за счет выбора достаточно малых δ и $\bar{g}^n$, достаточно близких к $g(n\delta)$. Но это и означает, что множество $\mathcal{D}(F_0)$ плотно в $\mathcal{B}'$. Поэтому для любых $u_0 \in \mathcal{B}$ и $g(\cdot) \in \mathcal{B}'$ определим *обобщенное решение* задачи (3.15) формулой (3.16), обосновывая это, как и в случае однородных задач, тем, что оно является пределом аппроксимирующих его подлинных решений и единственно, так как F есть единственное ограниченное расширение преобразования F_0 .

Система конечноразностных уравнений для задачи (3.15) имеет вид

$$B_1 u^{n+1} - B_0 u^n = g^{n+1}, \quad u^0 = u_0, \quad (3.19)$$

где B_1 и B_0 — разностные операторы, рассмотренные в § 3.2, а g^n или $g_{\Delta t}^n$ аппроксимируют $g(n\Delta t)$ или $g((n-1/2)\Delta t)$. Мы предположим просто, что $\|g_{\Delta t}^n - g(n\Delta t)\|$ может быть сделано произвольно малым равномерно по n при $0 \leq n\Delta t \leq T$ за счет подходящего выбора Δt . Как и в § 3.2, эти разностные уравнения можно переписать в виде

$$u^{n+1} - C(\Delta t) u^n = D(\Delta t) g^{n+1},$$

где операторы $C(\Delta t) = B_1^{-1} B_0$ и $D(\Delta t) = B_1^{-1}$ ограничены и зависят только от Δt , ибо по предположению все Δx , Δy и т. д. выражаются через Δt . Для явной схемы $D(\Delta t)$ есть просто $\Delta t I$, где I — тождественный оператор. Во всяком случае мы считаем, что элементы $\bar{g}^n$, определяемые равенством $D(\Delta t) g^n = \Delta t \bar{g}^n$, аппроксимируют g^n равномерно по n при достаточно малых Δt .

После многократного применения рассматриваемой разностной схемы получим

$$u^n = C(\Delta t)^n u_0 + \Delta t \sum_{j=1}^n C(\Delta t)^{n-j} \bar{g}^j. \quad (3.20)$$

Покажем, что если выполнены условия теоремы Лакса, то сумма в (3.20) аппроксимирует интеграл из (3.16), откуда и будет следовать сходимость приближенного решения к точному. Из равномерной непрерывности $g(s)$ и равномерной ограниченности операторов $E(t)$ и $C(\Delta t)$ (см. (3.9)) следует, что если при фиксированном t' последовательности значений Δt и m выбираются таким образом, что $\Delta t \rightarrow 0$ и $m\Delta t \rightarrow t'$, то величина

$$\| [E(m\Delta t) - C(\Delta t)^m] g(s) \|$$

стремится к нулю равномерно относительно s для $0 \leq s \leq T$. (Мы уже показали в § 3.5, что эта сходимость равномерна по t' при фиксированном s .) Поэтому для данного $\varepsilon > 0$ можно подобрать столь малое Δt , что интегральная сумма вида

$$\Delta t \sum_{j=0}^{n-1} E(t - j\Delta t) g(j\Delta t),$$

где $t = n\Delta t$, будет аппроксимировать с точностью ε интеграл из (3.16) и в свою очередь с той же точностью будет аппроксимироваться суммой

$$\Delta t \sum_{j=0}^{n-1} C(\Delta t)^{n-j} g(j\Delta t),$$

которую с точностью ε можно уже аппроксимировать суммой из равенства (3.20). Таким образом, сходимость доказана.

Отсюда следует, что теорема Лакса об эквивалентности справедлива также и для неоднородных краевых задач.

В дальнейшем мы будем обычно считать граничные условия однородными, так как от неоднородности можно избавиться, сузив область определения оператора A . Действительно, рассмотрим краевую задачу с неоднородными граничными условиями.

$$\begin{aligned} \frac{d}{dt} u(t) - Au(t) &= g(t), \\ A_1 u(t) &= h(t), \quad u(0) = u_0. \end{aligned}$$

Во втором уравнении этой системы, задающей граничные условия, A_1 — некоторый линейный оператор, а $h(t)$ — заданная функция; этот оператор выводит нас из банахова пространства $\mathcal{B}$ в некоторое другое пространство функций, определенных

лишь на границе, и $h(t)$ является именно такой функцией. Пусть $w(t)$ — некоторая гладкая функция из $\mathcal{B}$, значения которой принадлежат областям определения операторов A и A_1 , такая, что $A_1 w(t) = h(t)$; функцию $w(t)$ можно выбирать с большим произволом, поскольку она должна удовлетворять только граничным условиям. Теперь нашу задачу можно свести к предыдущей следующим хорошо известным способом. Пусть функция $v(t) = u(t) - w(t)$ удовлетворяет системе уравнений

$$\frac{d}{dt} v(t) - Av(t) = g(t) - \frac{d}{dt} w(t) - Aw(t),$$

$$A_1 v(t) = 0, \quad v(0) = u_0 - w(0),$$

правые части которых известны. Граничные условия здесь снова однородны, так что если теперь заменить оператор A его сужением A_0 на такие элементы v из $\mathcal{B}$, для которых $A_1 v = 0$, то мы получим задачу типа (3.15), к которой применимы все предыдущие рассуждения этого параграфа.

Р. Дж. Томпсон [1964] распространил большую часть приведенных здесь результатов на квазилинейный случай, в котором $g(t)$ заменяется на $g(t, \bar{u}(t))$, где $g(t, u)$ — функция, определенная при $0 \leq t \leq T$ и $u \in \mathcal{B}$, непрерывная по t при каждом u и удовлетворяющая равномерно по t условию Липшица относительно u в том смысле, что $\|g(t, u) - g(t, v)\| \leq M\|u - v\|$ при любых u, v и t (M — константа).

§ 3.8. Изменение нормы

Хотя выбор подходящей нормы в пространстве $\mathcal{B}$ является частью постановки задачи, и задача может быть корректно поставленной в одной норме и не обладать этим свойством в другой, тем не менее при выборе нормы имеется много произвола.

Предположим, что наряду с нормой $\|u\|$ для всех $u \in \mathcal{B}$ определена еще одна неотрицательная функция $\|u\|'$, обладающая всеми свойствами нормы, указанными во второй главе. Нормы $\|u\|$ и $\|u\|'$ называются *эквивалентными*, если существуют такие положительные постоянные K_1 и K_2 , что для любого $u \in \mathcal{B}$

$$K_1 \|u\| \leq \|u\|' \leq K_2 \|u\|. \quad (3.21)$$

Все топологические свойства пространства остаются неизменными при замене данной нормы эквивалентной ей нормой. В частности, свойства сходимости последовательности элементов, ограниченности множества элементов, ограниченности линейного оператора и равномерной ограниченности семейства операторов инвариантны относительно такой замены.

§ 3.9. Устойчивость и возмущения

Следующая теорема (Крайс [1962], Стрэнг [1964]) показывает, что устойчивость не нарушается малыми возмущениями.

Если разностная схема

$$u^{n+1} = C(\Delta t) u^n \quad (3.22)$$

устойчива, а семейство операторов $Q(\Delta t)$ ограничено, то разностная схема

$$u^{n+1} = [C(\Delta t) + \Delta t Q(\Delta t)] u^n \quad (3.23)$$

также устойчива.

По определению, устойчивость схемы (3.22) означает существование такой константы K , что

$$\|C(\Delta t)^n\| < K \quad \text{при} \quad \begin{cases} 0 < \Delta t < \tau, \\ 0 \leq n \Delta t \leq T. \end{cases} \quad (3.24)$$

Согласно условию теоремы существует такая константа H , что

$$\|Q(\Delta t)\| \leq H \quad \text{при} \quad 0 < \Delta t < \tau. \quad (3.25)$$

Выражение $(C + \Delta t Q)^n$ содержит 2^n слагаемых, причем для каждого $j = 0, 1, 2, \dots, n$ имеется $\binom{n}{j}$ слагаемых, состоящих из j множителей $\Delta t Q$ и $n - j$ множителей C , расположенных в определенном порядке. В каждом таком слагаемом множители C образуют не более чем $j + 1$ групп, поскольку их разделяет только j множителей $\Delta t Q$, и норма каждой группы не превосходит K . Поэтому

$$\begin{aligned} \|(C + \Delta t Q)^n\| &\leq \sum_{j=0}^n \binom{n}{j} K^{j+1} (\Delta t H)^j = \\ &= K(1 + \Delta t K H)^n \leq K e^{n \Delta t K H} \leq K e^{T K H}, \end{aligned} \quad (3.26)$$

откуда и следует устойчивость возмущенной системы.

Стоит отметить, что при некоторых других определениях устойчивости она более чувствительна к возмущениям. В одном из таких определений, которого придерживается ряд авторов, вместо равномерной ограниченности множества операторов (3.24) требуется лишь, чтобы они росли по норме не быстрее некоторой степени $1/\Delta t$ при $\Delta t \rightarrow 0$. Тогда, как было показано Крайсом на различных примерах, один из которых приведен в § 5.2, совсем простое возмущение может привести к экспоненциальному росту норм соответствующего множества операторов, причем в показатель экспоненты будет входить множитель $(1/\Delta t)^{1-1/p}$, где p — число компонент вектора-решения u .

Если краевая задача поставлена некорректно в том смысле, что хотя область определения $E_0(t)$ и плотна в $\mathcal{B}$, но семейство $E_0(t)$ не ограничено ни на одном интервале $0 \leq t < \tau$, то не существует разностной схемы, которая была бы согласована с этой задачей и в то же время устойчива. Действительно, из неограниченности $E_0(t)$ следует, что существует такая последовательность точных решений $u_j(t) = E_0(t) u_j^0$ ($j = 1, 2, \dots$), нормированных, скажем, условием $\|u_j^0\| = 1$ для каждого j , и такая сходящаяся к нулю последовательность $t_j > 0$, что

$$\|u_j(t_j)\| = \|E_0(t_j) u_j^0\| \rightarrow \infty \quad \text{при} \quad j \rightarrow \infty. \quad (3.27)$$

Пусть задано $\varepsilon > 0$. Тогда для любого точного решения $u_j(t)$ существует согласно условию согласованности (3.6с) такое $\Delta_j t = t/n_j$ (где n_j — некоторое натуральное число), что $\|[C(\Delta_j t) - E(\Delta_j t)]u_j(t)\| < \varepsilon \Delta_j t$. Если допустить теперь, что схема устойчива, то это приведет к противоречию. В самом деле из (3.10) и (3.11) следует, что

$$\|[C(\Delta_j t)^{n_j} - E(t_j)]u_j^0\| < \varepsilon K_2 T,$$

где K_2 определяется условием устойчивости, так что $\|C(\Delta_j t)^{n_j}\| \geq \|C(\Delta_j t)^{n_j} u_j^0\| \rightarrow \infty$ в силу (3.27) и того факта, что $\|u_j^0\| = 1$.

Таким образом, бесполезно искать устойчивые разностные схемы для таких, например, краевых задач, как задача Коши для уравнения Лапласа $\partial^2 u / \partial t^2 + \partial^2 u / \partial x^2 = 0$ или для уравнения теплопроводности с обратным ходом времени¹⁾.

¹⁾ Приведенные задачи, равно как и многие другие некорректные задачи, могут быть численно решены разностным методом с применением надлежащей регуляризации, например по Тихонову. — *Прим. ред.*

Глава 4

ЛИНЕЙНЫЕ КРАЕВЫЕ ЗАДАЧИ С ПОСТОЯННЫМИ КОЭФФИЦИЕНТАМИ

§ 4.1. Класс задач

В этой главе мы будем рассматривать линейные задачи с постоянными коэффициентами и с простыми граничными условиями, так что для представления их решений можно будет применять ряды или интегралы Фурье. В этом случае, как будет видно, понятия, введенные в предыдущей главе, принимают особенно простой вид. В частности, преобразование Фурье приводит условие устойчивости к такой форме, которая дает возможность получить целый ряд пригодных для практики признаков устойчивости. Простейшим из них и исторически первым является необходимое условие фон Неймана; в первом издании этой книги было приведено несколько достаточных условий для специальных случаев. Благодаря работам Като [1960], Крайса [1962] и Баченана [1963a, b] они были существенно улучшены и в конце концов сведены к одному простому алгебраическому условию устойчивости, которое позволяет в принципе решить вопрос об устойчивости любой разностной схемы с постоянными коэффициентами; получение этого общего условия является одним из основных вопросов данной главы. Будет рассмотрен ряд частных случаев с целью упростить исследование устойчивости на практике. Работа Крайса положила начало другому направлению в вопросе исследования устойчивости конкретных схем, так называемому энергетическому методу, с которым мы познакомимся подробнее в шестой главе.

Чтобы иметь возможность применять метод преобразования Фурье, мы предположим, что либо граничные условия могут быть заменены условиями периодичности, как в случае простейшей задачи теплопроводности в первой главе, либо рассматриваемые функции интегрируемы с квадратом во всем пространстве, как в квантовой механике. Для представления функций с помощью рядов или интегралов Фурье используются соответственно теоремы Рисса — Фишера и Планшереля, в которых фигурирует норма пространства L_2 . Поэтому банахово пространство $\mathcal{B}$ определяется однозначно выбором функциональных элементов u . Последовательность коэффициентов Фурье или результат преобразования Фурье задают точку во втором банаховом пространстве $\mathcal{B}'$, причем это соответствие является

взаимно однозначным и сохраняет норму, так что пространства $\mathcal{B}$ и $\mathcal{B}'$ изоморфны между собой. Поэтому все рассуждения предыдущей главы применимы к пространству $\mathcal{B}'$ точно так же, как и к пространству $\mathcal{B}$. Это весьма полезное обстоятельство обязано специальному выбору нормы пространства L_2 и не имеет места при другом выборе нормы, например в том случае, когда используется норма пространства C (т. е. максимум модуля функции), которая иногда считается наиболее естественной нормой численного анализа.

Указанные банаховы пространства $\mathcal{B}$ и $\mathcal{B}'$ являются гильбертовыми пространствами. Каждой паре элементов u и v из пространства $\mathcal{B}$ сопоставлено комплексное число, обозначаемое через (u, v) и называемое *скалярным произведением* u и v . Если u, v и w — произвольные элементы из $\mathcal{B}$, а α и β — комплексные числа, то

$$(u, \alpha v + \beta w) = \alpha (u, v) + \beta (u, w),$$

$$(v, u) = \overline{(u, v)},$$

$$(u, u) \geq 0,$$

$$(u, u) = 0 \text{ тогда и только тогда, когда } u = 0.$$

Первое равенство показывает, что функция (u, v) линейна относительно второго аргумента; в сочетании со вторым равенством это значит, что она антилинейна по первому аргументу в том смысле, что $(\alpha u + \beta v, w) = \bar{\alpha} (u, w) + \bar{\beta} (v, w)$, где черта сверху обозначает комплексное сопряжение. Скалярное произведение связано с нормой соотношением $(u, u) = \|u\|^2$.

§ 4.2. Ряды и интегралы Фурье

Если имеется p функций от d пространственных переменных, то для функций, соответствующих данной точке банахова пространства, используется векторное обозначение $u(x)$, где u — вектор с p компонентами, а x — вектор с d компонентами. Решение краевой задачи будет обозначаться через $u(x, t)$.

Если вектор-функция $u(x)$ периодична по каждому аргументу x_i с периодом L_i , то ее ряд Фурье имеет вид

$$u(x) = \left(\frac{1}{L_1 L_2 \dots L_d} \right)^{1/2} \sum_{k \in \mathcal{L}} \hat{u}(k) e^{-ik \cdot x}. \quad (4.1a)$$

Здесь k есть d -мерный вектор, который при суммировании пробегает решетку $\mathcal{L}$, состоящую из векторов, компоненты k_i которых независимо друг от друга принимают всевозможные значения вида $2\pi r_i / L_i$, где r_i — произвольное целое число. Векторный коэффициент Фурье $\hat{u}(k)$, рассматриваемый как функция аргу-

мента $\mathbf{k}$ на множестве $\mathcal{L}$, определяет точку в банаховом пространстве $\mathcal{B}'$. Обращение соотношения (4.1a) таково:

$$\hat{\mathbf{u}}(\mathbf{k}) = \left(\frac{1}{L_1 L_2 \dots L_d} \right)^{1/2} \int_0^L \mathbf{u}(\mathbf{x}) e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x}, \quad (4.2a)$$

где для сокращения записи введено обозначение

$$\int_0^L d\mathbf{x} = \int_0^{L_1} dx_1 \dots \int_0^{L_d} dx_d.$$

Имеет место так называемое равенство Парсевала

$$\int_0^L |\mathbf{u}(\mathbf{x})|^2 d\mathbf{x} = \sum_{\mathbf{k} \in \mathcal{L}} |\hat{\mathbf{u}}(\mathbf{k})|^2, \quad (4.3a)$$

где

$$|\mathbf{u}|^2 = \sum_{j=1}^p |u_j|^2;$$

для функции $\mathbf{u}(\mathbf{x})$ из $\mathcal{B}$ норма $\mathbf{u}$ определяется как корень квадратный из левой части равенства (4.3a), а для функции $\hat{\mathbf{u}}(\mathbf{k})$ из $\mathcal{B}'$ — как корень квадратный из правой его части. Таким образом, соотношения (4.1a) и (4.2a) устанавливают сохраняющее норму взаимно однозначное соответствие между элементами пространств $\mathcal{B}$ и $\mathcal{B}'$.

Если же вектор-функция $\mathbf{u}(\mathbf{x})$ интегрируема с квадратом во всем пространстве, то ее можно представить с помощью интеграла Фурье в следующем виде:

$$\mathbf{u}(\mathbf{x}) = \frac{1}{(2\pi)^{d/2}} \int \hat{\mathbf{u}}(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x}} d\mathbf{k}, \quad (4.1b)$$

где интегрирование теперь ведется по всему пространству переменных $k_1, k_2, \dots, k_d$, а заданная в этом пространстве функция $\hat{\mathbf{u}}(\mathbf{k})$ определяется равенством

$$\hat{\mathbf{u}}(\mathbf{k}) = \frac{1}{(2\pi)^{d/2}} \int \mathbf{u}(\mathbf{x}) e^{-i\mathbf{k} \cdot \mathbf{x}} d\mathbf{x}. \quad (4.2b)$$

В этом случае равенство Парсевала имеет вид

$$\int |\mathbf{u}(\mathbf{x})|^2 d\mathbf{x} = \int |\hat{\mathbf{u}}(\mathbf{k})|^2 d\mathbf{k}. \quad (4.3b)$$

Для достаточно хороших функций все эти равенства можно понимать буквально, но если запас элементов $\mathcal{B}$ и $\mathcal{B}'$ ограничить только такими функциями, то эти пространства не будут полными. Поэтому для соблюдения необходимого уровня строгости

следует принять все положения теории пространства L_2 и, в частности, считать две функции равными, если они отличаются друг от друга лишь на множестве меры нуль, интегралы понимать в смысле Лебега, а пределы в (4.1a), (4.1b) и в (4.2b) — в смысле сходимости в среднем, т. е. по норме соответствующего банахова пространства. Так, если $u_n(x)$ — последовательность частичных сумм ряда (4.1a), для которых, например, каждое r_i изменяется от $-n$ до n , то равенство (4.1a) понимается в том смысле, что

$$\|u - u_n\| \rightarrow 0 \quad \text{при} \quad n \rightarrow \infty.$$

Аналогично, если $u_k(x)$ есть интеграл типа (4.1b), для которого каждая переменная интегрирования k_i пробегает интервал $-K \leq k_i \leq K$, то равенство (4.1b) означает, что

$$\|u - u_k\| \rightarrow 0 \quad \text{при} \quad K \rightarrow \infty.$$

Теперь пространства $\mathcal{B}$ и $\mathcal{B}'$ полны. Преобразование (4.2a) отображает все пространство $\mathcal{B}$ в $\mathcal{B}'$ и в силу теоремы Фурье является обратным по отношению к (4.1a). По теореме Рисса — Фишера преобразование (4.1a) переводит все пространство $\mathcal{B}'$ в $\mathcal{B}$, так что любая определенная на $\mathcal{L}$ последовательность $\hat{u}(k)$, для которой сходится сумма из правой части равенства (4.3a), представляет собой последовательность коэффициентов Фурье некоторой функции $u(x) \in \mathcal{B}$. Аналогично (4.1b) в силу теоремы Планшереля является взаимно однозначным отображением пространства $\mathcal{B}'$ на пространство $\mathcal{B}$, а (4.2b) — обратным к нему.

В дальнейшем все это будет неявно предполагаться, если применяются соотношения типа (4.1a) — (4.3b). Обычно в обоих случаях мы будем использовать обозначения (4.1b) — (4.3b), подразумевая под $\mathcal{L}$ либо все пространство переменных $k_1, k_2, \dots, k_d$, либо описанную выше бесконечную дискретную решетку.

§ 4.3. Корректно поставленные краевые задачи

Общий линейный дифференциальный оператор с постоянными коэффициентами, действующий на функцию из $\mathcal{B}$, можно формально определить так: надо взять функцию $P(q)$, или $P(q_1, \dots, q_d)$, являющуюся $p \times p$ -матрицей, элементы которой суть полиномы относительно $q_1, q_2, \dots, q_d$, и затем заменить q_1 на $\partial/\partial x_1$, q_2 на $\partial/\partial x_2$ и т. д. Если A — такой оператор и мы применяем его к элементу $\hat{u}e^{ik \cdot x}$, где $\hat{u}$ — постоянный вектор, то результатом будет произведение этого элемента на матрицу $P(ik)$. Следовательно, если решение краевой задачи выражается интегралом Фурье

$$u(x, t) = (2\pi)^{-d/2} \int \hat{u}(k, t) e^{ik \cdot x} dk,$$

то преобразование Фурье $\hat{u}(k, t)$ должно удовлетворять уравнению

$$\frac{\partial}{\partial t} \hat{u}(k, t) = P(ik) \hat{u}(k, t),$$

которое служит эквивалентом в пространстве $\mathcal{B}'$ исходного дифференциального уравнения

$$\frac{\partial}{\partial t} u(x, t) = Au(x, t).$$

Поэтому

$$\hat{u}(k, t) = e^{iP(ik)t} \hat{u}(k, 0),$$

где $\hat{u}(k, 0)$ — преобразование Фурье для начального элемента $u(x, 0)$. Возвращаясь к $\mathcal{B}$, получаем

$$u(x, t) = (2\pi)^{-d/2} \int e^{iP(ik)t} \hat{u}(k, 0) e^{ik \cdot x} dk, \quad (4.4)$$

где

$$\hat{u}(k, 0) = (2\pi)^{-d/2} \int u(x, 0) e^{-ik \cdot x} dx.$$

Известно, что функция e^M от $p \times p$ -матрицы M определяется степенным рядом

$$e^M = 1 + M + \frac{1}{2!} M^2 + \frac{1}{3!} M^3 + \dots,$$

который поточечно абсолютно сходится и обладает тем свойством, что

$$\frac{d}{dt} e^{tM} = M e^{tM}.$$

Первое требование корректной постановки краевых задач, а именно что область определения разрешающего оператора $E_0(t)$ плотна в $\mathcal{B}$ (см. гл. 3), автоматически выполняется для рассматриваемых в этой главе задач. Это объясняется тем, что решения такого рода законны всякий раз, когда суммирование или интегрирование в (4.4) является конечным, т. е. когда начальный элемент $u(x, 0)$ есть тригонометрический полином или же $u(x, 0)$ имеет преобразование Фурье с компактным носителем¹⁾, но такие элементы в обоих случаях плотны в $\mathcal{B}$. Второе требование корректной постановки краевых задач состоит в том, что $\|e^{tP(ik)}\|$ должны быть равномерно ограничены по k ; тогда

$$\|E(t)\| = \max_{k \in \mathcal{L}} \|e^{tP(ik)}\|. \quad (4.5)$$

¹⁾ Носителем функции называется замыкание множества тех точек, в которых эта функция отлична от нуля. Говорят, что функция имеет компактный носитель, если ее носитель является компактным множеством.

Выполнение этого условия должно проверяться отдельно в каждом данном случае. В качестве простого примера рассмотрим уравнение теплопроводности, для которого $A = \sigma \partial^2 / \partial x^2$. Тогда величина $\|e^{tp(ik)}\| = \exp(-\sigma tk^2)$ ограничена единицей, если $\sigma t \geq 0$. Но если $\sigma t < 0$, то ее можно сделать сколь угодно большой за счет выбора достаточно большого k ; поэтому задача теплопроводности, решаемая назад по времени, не является корректно поставленной.

Однако цель этой книги — в изучении решений задач, а не их постановок, поэтому мы будем включать свойство корректности в постановку краевых задач, другими словами, мы будем предполагать, что нам заданы только корректно поставленные задачи. Следует подчеркнуть, что в таком случае выбор банахова пространства является частью постановки задачи: как мы видели выше, выбор среднеквадратичной нормы определяется желанием использовать аппарат преобразования Фурье, но для самого вектора u , норма которого определена таким образом, еще часто остается некоторая свобода выбора. Это, в частности, относится к тому случаю, когда перед применением общей теории дифференциальное уравнение высокого порядка сводится к системе уравнений первого порядка. Чтобы проиллюстрировать это, мы покажем в гл. 10, что задача о волновом движении является корректно поставленной при одном простом выборе банахова пространства и не является таковой при другом столь же простом его выборе. В этих простейших случаях проверка корректности обычно сравнительно несложна; в других случаях, вроде задачи о многомерном движении жидких сред, в настоящее время не ясно, как ставить корректные задачи, не привлекая дополнительных физических принципов.

§ 4.4. Конечноразностные уравнения

Мы предполагаем, что разностные уравнения даны в виде двухслойных формул. Многослойные формулы будут рассмотрены в гл. 7.

Пусть $u^n = u^n(x)$ — приближение для $u(x, n\Delta t)$, полученное из системы разностных уравнений. Обозначим через T^β оператор, заменяющий значение функции в точке $x = (x_1, x_2, \dots, x_d)$ ее значением в соседней точке $(x_1 + \beta_1 \Delta x_1, x_2 + \beta_2 \Delta x_2, \dots, x_d + \beta_d \Delta x_d)$ пространственной сетки, где компоненты $\beta_1, \beta_2, \dots, \beta_d$ векторного индекса β принимают целые значения. Каждое разностное уравнение в точке x обращает в нуль определенную линейную комбинацию компонент векторов u^n и u^{n+1} на некоторой группе таких соседних точек. Таким образом, систему разностных уравнений

$$B_1 u^{n+1} - B_0 u^n = 0$$

можно записать в виде

$$\sum_{\beta \in N_1} B_1^\beta T^\beta \mathbf{u}^{n+1} - \sum_{\beta \in N_0} B_0^\beta T^\beta \mathbf{u}^n = 0, \quad (4.6)$$

где суммирование ведется по конечным множествам N_1 и N_0 точек, соседних¹⁾ с точкой $\mathbf{x}$. Для каждого β B_0^β и B_1^β суть $p \times p$ -матрицы с постоянными элементами; последнее означает, что элементы этих матриц могут зависеть от Δt , Δx_1 и т. д., но не могут зависеть от $\mathbf{x}$ или от t . Для явной системы разностных уравнений множество N_1 сводится к единственной точке $\beta = 0$; тогда матрица B_1^0 обычно диагональна и во всяком случае невырождена. Если уравнения неявные, то мы предположим, что они всегда разрешимы для периодического $\mathbf{u}^n$ и что решение $\mathbf{u}^{n+1}$ однозначно определяется уравнениями и условием периодичности.

При вычислениях уравнения (4.6) используются только в точках сетки, но их с таким же успехом можно применять и в других точках, так что если вектор $\mathbf{u}^n(\mathbf{x})$ задан для всех $\mathbf{x} = (x_1, \dots, x_d)$, то $\mathbf{u}^{n+1}(\mathbf{x})$ определяется для всех $\mathbf{x}$ уравнением (4.5) и условием периодичности.

Если в уравнения (4.6) подставить интегралы Фурье для $\mathbf{u}^n(\mathbf{x})$ и $\mathbf{u}^{n+1}(\mathbf{x})$, то типичный член будет содержать множитель

$$\exp \{i [k_1 (x_1 + \beta_1 \Delta x_1) + \dots + k_d (x_d + \beta_d \Delta x_d)]\},$$

и мы можем сократить на величину $\exp \{ik \cdot \mathbf{x}\}$ все члены уравнения. То, что останется после сокращения, можно записать в виде

$$H_1 \mathbf{v}^{n+1}(\mathbf{k}) - H_0 \mathbf{v}^n(\mathbf{k}) = 0, \quad (4.7)$$

где H_1 обозначает матрицу

$$H_1 = \sum_{\beta \in N_1} B_1^\beta \exp \{i [k_1 \beta_1 \Delta x_1 + \dots + k_d \beta_d \Delta x_d]\}, \quad (4.8)$$

а H_0 определяется аналогично. Как и в предыдущей главе, мы предположим, что способ измельчения ячеек сетки должен быть установлен заранее, так что $\Delta x_1, \dots, \Delta x_d$ стремятся к нулю вместе с Δt по определенным законам

$$\Delta x_1 = g_1(\Delta t) \quad \text{и т. д.} \quad (4.9)$$

При этих условиях матрицы H_1 и H_0 зависят только от Δt и $\mathbf{k}$. Предположение о разрешимости, сделанное в предыдущем абзаце, равносильно требованию, что матрица H_1 имеет обратную. Следовательно, можно написать

$$\mathbf{v}^{n+1}(\mathbf{k}) = G(\Delta t, \mathbf{k}) \mathbf{v}^n(\mathbf{k}), \quad (4.10)$$

¹⁾ Сама точка $(x_1, \dots, x_d)$, конечно, также вообще говоря включается в множество N_1 .

где

$$G(\Delta t, k) = H_1^{-1} H_0. \quad (4.11)$$

Ясно, что $G(\Delta t, k)$ — это представление в пространстве $\mathcal{B}'$ разностного оператора $C(\Delta t)$, фигурирующего в теореме Лакса, а уравнение (4.10) — это не что иное, как форма, которую принимает уравнение (3.5) при отображении в пространство преобразований Фурье $\mathcal{B}'$. Матрица $G(\Delta t, k)$ называется *матрицей перехода*¹⁾.

§ 4.5. Порядок точности и условие согласованности

В уравнении (4.6) имеется очевидная неопределенность, состоящая в том, что оно может быть умножено на произвольную невырожденную матрицу, и в результате получится уравнение, эквивалентное исходному. В случае явной системы уравнений этим произволом можно иногда воспользоваться для того, чтобы представить ее в виде, разрешенном относительно u^{n+1} (например, при решении простейшей задачи теплопроводности умножение конечноразностного уравнения на Δt приводит к замене отношения $(u_j^{n+1} - u_j^n)/\Delta t$ разностью $u_j^{n+1} - u_j^n$). Однако в настоящей главе мы должны избегать этой неопределенности; будем предполагать, что левая часть уравнения (4.6) получается непосредственно из выражения $\partial u / \partial t - Au$ в результате замены производных разделенными разностями, а других величин — их значениями в узлах, полученными путем интерполяции или экстраполяции. Это означает, что разностный оператор в левой части уравнения (4.6) при применении к любой гладкой функции должен давать приближенно тот же самый эффект, что и оператор $\partial / \partial t - A$. Как и прежде, способ измельчения ячеек сетки предполагается заранее фиксированным соотношениями типа (4.9)²⁾. Определим *порядок точности* разностной системы как наибольшее действительное число p , для которого

$$\|B_1 u(t + \Delta t) - B_0 u(t)\| = O[(\Delta t)^p] \quad \text{при} \quad \Delta t \rightarrow 0 \quad (4.12)$$

для всех достаточно гладких точных решений дифференциального уравнения $\partial u / \partial t = Au$. Согласно нашему предположению о разрешимости разностной системы относительно u^{n+1} оператор B_1 должен иметь обратный, B_1^{-1} . Исходя из того что разностная схема получена непосредственно из дифференциаль-

¹⁾ В случае $p = 1$, когда матрица перехода состоит из единственного элемента, $G(\Delta t, k)$ называют также *множителем перехода*.

²⁾ На практике формулы (4.9) утверждают обычно, что Δx_i пропорциональны некоторым положительным рациональным степеням Δt . Тогда p легко находится и также является положительным рациональным числом.

ного уравнения, нетрудно убедиться, что главный член в разложении B_1 по возрастающим (возможно, дробным) степеням Δt пропорционален $(\Delta t)^{-1}$. Таким образом, $\|B_1^{-1}\| = O(\Delta t)$ (см. § 7.4, где более подробно рассматривается общая многослойная разностная схема), и потому соотношение (4.12) можно переписать в виде

$$\|u(t + \Delta t) - (B_1)^{-1} B_0 u(t)\| = O[(\Delta t)^{p+1}] \quad \text{при } \Delta t \rightarrow 0. \quad (4.12a)$$

Сравнивая (4.12a) с условием согласованности (3.6с) и учитывая, что $C = B_1^{-1} B_0$, мы видим, что условие согласованности и положительность порядка точности разностной системы эквивалентны между собой. Очень часто p является натуральным числом; в этом случае согласованность означает, что разностные уравнения имеют по меньшей мере первый порядок точности.

Применяя преобразование Фурье к выражению, стоящему под знаком нормы в (4.12a), можно дать двойственное определение порядка точности, использующее матрицу перехода G . Действительно, из (4.4), (4.7) и (4.8) следует, что

$$\|e^{\Delta t P(i\mathbf{k})} - G(\Delta t, \mathbf{k})\| = O[(\Delta t)^{p+1}] \quad \text{при } \Delta t \rightarrow 0. \quad (4.13)$$

Выражение под знаком нормы в (4.12) было названо в первой главе погрешностью аппроксимации. Ее легко можно оценить, представив функцию $u(\mathbf{x}, t)$ в виде ряда Тэйлора с остаточным членом. Гладкость функции u , которую мы предполагали выше, фактически определяется числом производных, необходимых для такого представления. Мы будем предполагать, что решения с нужной нам гладкостью плотны в классе всех точных решений дифференциального уравнения.

§ 4.6. Устойчивость

Мы уже указывали в § 4.5, что уравнение (4.10) соответствует уравнению $u^{n+1} = C(\Delta t)u^n$ общей теории, изложенной в гл. 3, и что это уравнение, следовательно, определяет некоторый оператор в $\mathcal{B}'$, эквивалентный оператору $C(\Delta t)$. Этот оператор сводится к умножению вектора v^n на матрицу $G(\Delta t, \mathbf{k})$ (куда $\mathbf{k}$ входит только как параметр). Очевидно поэтому, что условие устойчивости касается матриц

$$G(\Delta t, \mathbf{k})^n \quad \text{при} \quad \begin{cases} 0 < \Delta t < \tau, \\ 0 \leq n \Delta t \leq T \end{cases}$$

и фактически означает, что операторы, отвечающие этим матрицам, равномерно ограничены.

Если L — оператор в $\mathcal{B}'$, задаваемый матрицей $F(k)$, т. е. равенством $w(k) = F(k)v(k)$, то

$$\|L\| = \max_{k \in \mathcal{L}} \|F(k)\|, \quad (4.14)$$

где при заданном k норма матрицы $F(k)$ определяется формулой

$$\|F(k)\| = \max_{|v|=1} |F(k)v| = \max_{v \neq 0} \frac{|F(k)v|}{|v|}. \quad (4.15)$$

Здесь, как и в § 4.2, под длиной $|v|$ p -мерного вектора v подразумевается $[|v_1|^2 + \dots + |v_p|^2]^{1/2}$. Таким образом, $\|F(k)\|$ есть норма матрицы $F(k)$, рассматриваемой как преобразование p -мерного евклидова пространства векторов $v, w, \dots$.

Условие устойчивости означает, что для некоторого положительного τ матрицы

$$G(\Delta t, k)^n \quad \text{при} \quad \begin{cases} 0 < \Delta t < \tau, \\ 0 \leq n \Delta t \leq T \\ \text{и всех } k \in \mathcal{L} \end{cases} \quad (4.16)$$

равномерно ограничены.

Пусть F — матрица размера $p \times p$ и $\lambda_1, \dots, \lambda_p$ — ее собственные значения; тогда максимум $|\lambda_i|$ ($i = 1, \dots, p$) называется ее *спектральным радиусом* (или *радиусом спектра*). Если R — спектральный радиус матрицы F , то, очевидно, $\|F\| \geq R$, потому что максимум значений отношения $|Fv|/|v|$ не меньше, чем значение, получающееся при замене в этом отношении вектора v собственным вектором матрицы F , соответствующим наибольшему по модулю собственному значению. Спектральным радиусом матрицы F^n является R^n , так как собственными значениями матрицы F^n являются n -е степени собственных значений матрицы F . Кроме того,

$$\|F^2\| = \max_{v \neq 0} \frac{|F(Fv)|}{|Fv|} \frac{|Fv|}{|v|} \leq \max_{\substack{v \neq 0, \\ w=0}} \frac{|Fw|}{|w|} \frac{|Fv|}{|v|} = \|F\|^2$$

и аналогично $\|F^n\| \leq \|F\|^n$.

В дальнейшем мы будем обозначать через $R(\Delta t, k)$ спектральный радиус матрицы $G(\Delta t, k)$. В общем случае имеем¹⁾

$$R(\Delta t, k)^n \leq \|G(\Delta t, k)^n\| \leq \|G(\Delta t, k)\|^n. \quad (4.17)$$

¹⁾ Если матрица $G(\Delta t, k)$ нормальна, то соотношение (4.17) превращается в двойное равенство. Нормальной называется матрица, перестановочная со своей сопряженной (т. е. с эрмитово сопряженной). Для существования у матрицы полной системы ортогональных собственных векторов необходимо и достаточно, чтобы она была нормальна. Нормальная матрица может быть записана в виде $A + iB$, где A и B — перестановочные эрмитовы матрицы. Матрица перехода $G(\Delta t, k)$ весьма часто не является нормальной.

§ 4.7. Условие фон Неймана

Необходимым условием устойчивости является существование такой постоянной C_1 , что

$$R(\Delta t, k)^n \leq C_1 \quad \text{при} \quad \begin{cases} 0 < \Delta t < \tau, \\ 0 \leq n \Delta t \leq T \\ \text{и всех } k \in \mathcal{L}. \end{cases}$$

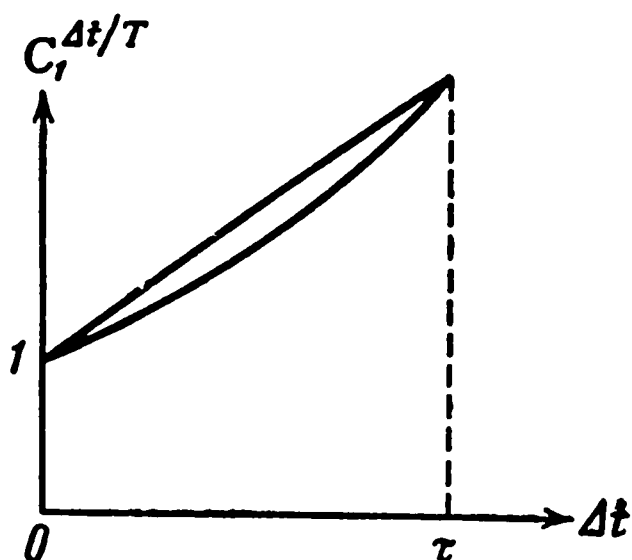
Не уменьшая общности, мы можем считать, что $C_1 \geq 1$. Следовательно,

$$R(\Delta t, k) \leq C_1^{1/n}, \quad 0 \leq n \leq \frac{T}{\Delta t},$$

и в частности

$$R(\Delta t, k) \leq C_1^{\Delta t/T}.$$

Для Δt из интервала $0 < \Delta t < \tau$ величина $C_1^{\Delta t/T}$ ограничена линейным выражением вида $1 + C_2 \Delta t$ (верхней линией на



Р и с. 4.1. Граница для части экспоненциальной кривой.

рис. 4.1). Вспоминая определение спектрального радиуса, получаем условие

$$|\lambda_i| \leq 1 + O(\Delta t) \quad \text{при} \quad \begin{cases} 0 < t < \tau, \\ i = 1, 2, \dots, p \\ \text{и всех } k \in \mathcal{L}, \end{cases} \quad (4.18)$$

где $\lambda_1, \lambda_2, \dots, \lambda_p$ суть собственные значения матрицы перехода $G(\Delta t, k)$. Условие (4.18) и является необходимым условием устойчивости фон Неймана.

В простейшей задаче теплопроводности (гл. 1) мы применяли условие устойчивости в виде

$$|\lambda_i| \leq 1. \quad (4.19)$$

Но в некоторых задачах бывает так, что компоненты точного решения с ростом времени возрастают экспоненциально, а условие (4.19) не допускает такого роста (и фактически совсем не может быть удовлетворено без нарушения согласованности).

В таких случаях нужно использовать условие (4.18), которое является более общим, чем (4.19), и допускает экспоненциальный рост компонент решения. В качестве примера рассмотрим модифицированную задачу теплопроводности

$$\frac{\partial u}{\partial t} = \sigma \frac{\partial^2 u}{\partial x^2} + bu, \quad \sigma > 0, \quad b > 0,$$

получающуюся из простейшей добавлением источников тепла. Для явной разностной схемы

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \sigma \frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2} + bu_j^n$$

множитель перехода имеет вид

$$G(\Delta t, k) = 1 - 4 \frac{\sigma \Delta t}{(\Delta x)^2} \sin^2 \frac{k \Delta x}{2} + b \Delta t,$$

и ясно, что условие Неймана (4.18) выполняется, если

$$\frac{\sigma \Delta t}{(\Delta x)^2} = \text{const} \leq 1/2,$$

как было установлено для немодифицированного уравнения.

Следует заметить, что множители перехода (собственные числа матрицы G) можно получить, если подставить $u_0 \exp(ik \cdot x)$ и $\lambda u_0 \exp(ik \cdot x)$ вместо u^n и u^{n+1} непосредственно в разностные уравнения (4.6); это даст p линейных уравнений относительно p компонент вектора u_0 и λ определяется условием равенства нулю определителя из коэффициентов этой системы. Такой подход может оказаться полезным для неявных разностных схем.

§ 4.8. Одно простое достаточное условие

В § 4.6 было отмечено, что для нормальной матрицы спектральный радиус равен ее норме. Следовательно, если матрица $G(\Delta t, k)$ нормальна, то условие фон Неймана является не только необходимым, но и достаточным условием устойчивости.

Это имеет место, в частности, в случае $p = 1$, потому что все матрицы первого порядка перестановочны. Мы приходим таким образом к выводу: для двухслойных разностных уравнений с одной зависимой переменной условие фон Неймана является не только необходимым, но и достаточным условием устойчивости. Это верно и для двухслойных разностных уравнений в задаче теплопроводности с произвольным числом независимых переменных.

В качестве дальнейшего примера упомянем разностные уравнения для одномерных звуковых волн, рассматриваемые в гл. 10.

Волновое уравнение имеет второй порядок по времени, так что $p = 2$. Для простейшего явного разностного уравнения матрица перехода равна

$$G = \begin{bmatrix} 1 & ia \\ ia & 1 - a^2 \end{bmatrix}, \quad (4.20)$$

где $a = (2c \Delta t / \Delta x) \sin(k \Delta x / 2)$; легко видеть, что G не является нормальной матрицей, потому что $G^* G - G G^* \neq 0$. Для простейшего неявного уравнения матрица перехода равна

$$G = \begin{bmatrix} \frac{1 - \frac{1}{4} a^2}{1 + \frac{1}{4} a^2} & \frac{ia}{1 + \frac{1}{4} a^2} \\ \frac{ia}{1 + \frac{1}{4} a^2} & \frac{1 - \frac{1}{4} a^2}{1 + \frac{1}{4} a^2} \end{bmatrix}$$

и, очевидно, унитарна, а значит нормальна. Поэтому условие фон Неймана достаточно для устойчивости такого неявного уравнения.

Несколько более общее достаточное условие можно получить из правого неравенства (4.17). А именно если для некоторого M и некоторого $\tau > 0$

$$\|G(\Delta t, k)\| \leq 1 + M \Delta t \quad \text{при} \quad 0 < \Delta t < \tau,$$

то разностные уравнения устойчивы, так как в этом случае

$$\|G\|^n \leq \exp(MT) \quad \text{при} \quad 0 \leq n \Delta t \leq T.$$

Во многих практических задачах ни одно из этих условий не выполняется. Поэтому необходимо иметь более совершенные критерии устойчивости.

§ 4.9. Теорема Крайса о матрицах

Прежде чем отыскивать дальнейшие условия, достаточные при определенных обстоятельствах для устойчивости, полезно сейчас изложить теорию устойчивых семейств матриц. Это позволит нам получить некоторые необходимые и достаточные критерии устойчивости, которые хотя и являются весьма громоздкими, но зато дают возможность без труда найти ряд достаточных условий. Кроме того, позже нам потребуется новое определение устойчивости, а именно такое: разностная система устойчива, если существует некоторая норма $\|\cdot\|_H$, эквивалентная данной и такая, что

$$\|u^{n+1}\|_H \leq [1 + O(\Delta t)] \|u^n\|_H,$$

т. е.

$$\|C(\Delta t)\|_H \leq 1 + O(\Delta t).$$

Один из результатов этого параграфа состоит в том, что для рассматриваемых задач с постоянными коэффициентами это новое определение эквивалентно принятому ранее.

Прежде всего удобно избавиться от ограничения на n , фигурирующего в (4.16). Назовем семейство $\mathcal{F}$ матриц A размера $p \times p$ устойчивым, если справедливо следующее утверждение:

(A) Существует такая постоянная C , что для любой матрицы $A \in \mathcal{F}$ и любого натурального числа ν

$$\|A^\nu\| \leq C. \quad (4.21)$$

Ясно, что если степени матрицы G из (4.16) равномерно ограничены числом K , то семейство матриц $\{e^{-\alpha \Delta t} G(\Delta t, k)\}$, где $\alpha = T^{-1} \ln K$, устойчиво в смысле данного определения. Действительно, любое натуральное число ν можно представить в виде¹⁾

$$\nu = m(T/\Delta t) + n, \quad \text{где } 0 \leq n \Delta t < T,$$

так что

$$\|(e^{-\alpha \Delta t} G)^\nu\| \leq \|(e^{-\alpha \Delta t} G)^{T/\Delta t}\|^m \|(e^{-\alpha \Delta t} G)^n\| \leq K.$$

Обратно, если для некоторой постоянной α семейство матриц $\{e^{-\alpha \Delta t} G(\Delta t, k)\}$ удовлетворяет условию (4.21), то при $0 \leq n \Delta t \leq T$

$$\|G^n\| \leq C e^{\alpha T} = \text{const.}$$

Следовательно, такое употребление термина «устойчиво» согласуется с уже изложенным. Далее, для справедливости утверждения (A) необходимо, чтобы все собственные значения λ_i матрицы A лежали внутри или на границе единичного круга. Это значит, что собственные значения λ_i матрицы G должны при некоторой постоянной α удовлетворять неравенству

$$|\lambda_i| \leq e^{\alpha \Delta t},$$

что эквивалентно условию фон Неймана (4.18).

Следующая теорема Крайса [1962] дает три важных критерия устойчивости семейства матриц.

Устойчивость семейства $\mathcal{F}$ матриц A эквивалентна каждому из следующих утверждений:

(R) Существует такая постоянная C_R , что для любой матрицы $A \in \mathcal{F}$ и любого комплексного числа z с $|z| > 1$ существует резольвента $(A - zI)^{-1}$

$$\|(A - zI)^{-1}\| \leq \frac{C_R}{|z| - 1}. \quad (4.22)$$

¹⁾ Очевидно, без ограничения общности можно считать, что T есть целое кратное Δt .

(S) Существуют такие постоянные C_S и C_B и для каждой матрицы $A \in \mathcal{F}$ такая невырожденная матрица S , что $\|S\|$, $\|S^{-1}\| \leq C_S$ и матрица $B = SAS^{-1}$ есть верхняя треугольная матрица, внедиагональные элементы которой удовлетворяют неравенству

$$|B_{ij}| \leq C_B \min\{1 - |\kappa_i|, 1 - |\kappa_j|\}, \quad (4.23)$$

где κ_i — диагональные элементы матрицы B , т. е. собственные значения как для B , так и для A (заметим, что отсюда следует, что $|\kappa_i| \leq 1$):

$$B = \begin{bmatrix} \kappa_1 & B_{12} & B_{13} & \dots & B_{1p} \\ 0 & \kappa_2 & B_{23} & \dots & B_{2p} \\ 0 & 0 & \kappa_3 & \dots & B_{3p} \\ \vdots & \vdots & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & & \kappa_p \end{bmatrix}.$$

(H) Существует такая постоянная $C_H > 0$ и для каждой матрицы $A \in \mathcal{F}$ такая положительно определенная эрмитова матрица H , что¹⁾

$$C_H^{-1}I \leq H \leq C_HI, \quad (4.24)$$

$$A^*HA \leq H. \quad (4.25)$$

Все известные для матриц общего вида достаточные условия устойчивости, формулируемые в виде некоторых ограничений на матричные элементы, основаны на приведении этих матриц к треугольному виду. При этом собственные значения представляются как диагональные элементы, и условие фон Неймана для них дополняется некоторыми неравенствами относительно внедиагональных элементов. Утверждение (S) служит примером условия такого типа; другие примеры появятся ниже.

Наше первое утверждение устанавливает верхнюю допустимую границу роста нормы резольвенты $(A - zI)^{-1}$ при $|z| \rightarrow 1$; поэтому это условие называется *резольвентным*. Оно оказывается чрезвычайно полезным в теоретических исследованиях и, в частности, будет для нас основным при доказательстве утверждения (S). Наконец, условие (H) обеспечивает существование новой нормы $\|\cdot\|_H$: в пространстве $\mathcal{B}'$ преобразований Фурье положим

$$\|v\|_H^2 = v^*Hv,$$

¹⁾ Говоря о двух матрицах A и H , что $A \leq H$, мы имеем в виду, что $u^*Au \leq u^*Hu$ для любого вектора u .

так что если $G = e^{\alpha \Delta t} A$, то в силу (4.25)

$$\|v^{n+1}\|_H^2 = e^{2\alpha \Delta t} (v^n)^* A^* H A v^n \leq [1 + O(\Delta t)] \|v^n\|_H^2.$$

Теперь нетрудно доказать справедливость аналогичного неравенства в исходном пространстве $\mathcal{B}$.

Другая формулировка этого утверждения состоит в том, что векторное пространство, в котором действует оператор A , можно преобразовать таким образом, что норма преобразованного оператора A будет ограничена единицей. Действительно, для каждой матрицы H можно найти такую матрицу T , что $T^* T = H$ и $\|T\|, \|T^{-1}\| \leq C_H^{1/2}$. Тогда в пространстве векторов $w = Tv$ матрица A заменится на TAT^{-1} , и

$$\begin{aligned} \|TAT^{-1}\|^2 &= \max_{|u|=1} u^* T^{-1*} A^* T^* T A T^{-1} u \leq \\ &\leq \max_{|u|=1} u^* T^{-1*} T^* T T^{-1} u = 1. \end{aligned}$$

Для доказательства теоремы Крайса мы покажем, что из условия (A) вытекает условие (R), из условия (R) — условие (S), из условия (S) — условие (H) и, наконец, что из условия (H) вытекает условие (A). Наиболее трудным является второй шаг, поэтому мы оставим его на самый конец.

Как мы уже отмечали, если выполнено (A), то собственные значения λ_i матрицы A по модулю не превосходят единицы, поэтому $(A - zI)^{-1}$ существует для $|z| > 1$. Кроме того,

$$\|(A - zI)^{-1}\| = \left\| \sum_{v=0}^{\infty} A^v z^{-v-1} \right\| \leq C \sum_{v=0}^{\infty} |z|^{-v-1} = C(|z| - 1)^{-1},$$

так что выполняется (R), причем $C_R \leq C$.

Чтобы вывести (H) из (S), введем вспомогательную диагональную матрицу

$$D = \begin{bmatrix} \Delta & & & 0 \\ & \Delta^2 & & \\ & & \ddots & \\ 0 & & & \Delta^p \end{bmatrix},$$

где $\Delta > 1$. Покажем, что при достаточно больших значениях Δ

$$D - B^* D B \geq 0, \quad (4.26)$$

т. е.

$$M = I - (D^{-1/2} B^* D^{1/2})(D^{1/2} B D^{-1/2}) \geq 0.$$

Для доказательства воспользуемся теоремой Гершгорина, которая утверждает, что если вокруг каждой точки M_{ii} ($i = 1, 2, \dots$

..., p) в комплексной плоскости описан круг радиуса $\sum_{j \neq i} |M_{ij}|$, то все собственные значения M лежат в объединении этих кругов¹⁾. Так как ij -й элемент матрицы $D^{1/2}BD^{-1/2}$ равен $\Delta^{(i-j)/2}B_{ij}$, то i -й диагональный элемент матрицы M равен

$$M_{ii} = 1 - \sum_{k=1}^n |B_{ki}|^2 \Delta^{k-i} = 1 - |\kappa_i|^2 + \varepsilon_i,$$

где $|\varepsilon_i| \leq C_B^2 (1 - |\kappa_i|)^2 / (\Delta - 1)$ в силу неравенства (4.23); сумма же внедиагональных элементов равна

$$\delta_i = \sum_{j \neq i} |M_{ij}| \leq \sum_{j \neq i} \sum_{k=1}^n |B_{ki}| |B_{kj}| \Delta^{(2k-i-j)/2}.$$

Поскольку матрица B треугольна, во внутренней сумме могут быть отличными от нуля только те слагаемые, для которых $k \leq i$ и $k \leq j$. Так как $i \neq j$, то $2k - i - j \leq -1$, и после несложных преобразований получаем

$$\delta_i \leq C_B^2 (1 - |\kappa_i|) 2(p-1) / (\Delta^{1/2} - 1).$$

Выберем теперь Δ так, чтобы $\delta_i + |\varepsilon_i| \leq 1 - |\kappa_i|^2$ для всех i . Это можно сделать, поскольку

$$\begin{aligned} \delta_i + |\varepsilon_i| &\leq C_B^2 \left[\frac{2(p-1)(1-|\kappa_i|)}{\Delta^{1/2}-1} + \frac{(1-|\kappa_i|)^2}{\Delta-1} \right] \leq \\ &\leq (1-|\kappa_i|^2) C_B^2 \left[\frac{2(p-1)}{\Delta^{1/2}-1} + \frac{1}{\Delta-1} \right] \end{aligned}$$

и коэффициент при $1 - |\kappa_i|^2$ не превосходит единицы для достаточно больших Δ . Так как матрица M эрмитова, то при указанном выборе Δ в силу теоремы Гершгорина все ее собственные значения неотрицательны, т. е. $M \geq 0$. Отсюда следует, что для таких Δ выполняется неравенство (4.26), что мы и хотели доказать. Таким образом,

$$S^{-1*} A^* S^* D S A S^{-1} - D \leq 0,$$

¹⁾ Пусть λ — собственное значение, а $v = (v_1, \dots, v_p)$ — соответствующий собственный вектор матрицы M , и пусть наибольшей по модулю компонентой v будет v_i . Тогда

$$(\lambda - M_{ii}) v_i = \sum_{j \neq i} M_{ij} v_j,$$

$$|\lambda - M_{ii}| |v_i| \leq \sum_{j \neq i} |M_{ij}| |v_j| \leq |v_i| \sum_{j \neq i} |M_{ij}|,$$

поэтому это собственное значение λ лежит в i -м круге, что и требовалось доказать. Ср. Гершгорин [1931].

т. е

$$A^*HA - H \leq 0,$$

где матрица $H = S^*DS$, очевидно, удовлетворяет всем требованиям утверждения (H). В частности, $C_H \leq \Delta^p C_s^2$.

Предположим теперь, что выполнено (H), и рассмотрим векторы $w_v = Aw_{v-1} = A^v w_0$. Имеем

$$w_v^* H w_v = w_{v-1}^* A^* H A w_{v-1} \leq w_{v-1}^* H w_{v-1} \leq \dots \leq w_0^* H w_0,$$

так что в силу (4.24) $|w_v|^2 \leq C_H^2 |w_0|^2$, т. е. $\|A^v\| \leq C_H$.

Нам осталось вывести (S) из (R). Доказательство будет основано на приводимой ниже лемме. Отметим прежде всего, что любую матрицу A можно привести к верхней треугольной форме U^*AU с помощью унитарного преобразования U ¹⁾. Так как резольвентное условие при этом не нарушается, можно считать, что каждая матрица A задана именно в такой форме.

Теперь займемся построением преобразования подобия S , которое позволит нам получить требуемый результат последовательно для каждой из верхних диагоналей матрицы A , т. е. установить неравенство (4.23) поочередно для значений $j - i = 1, 2, \dots, p - 1$.

В качестве первого шага заметим, что каждый угловой главный минор (левый верхний или правый нижний) матрицы A удовлетворяет условию (R) с той же константой C_R ²⁾. Далее, каждый элемент первой верхней диагонали матрицы A принадлежит некоторому минору второго порядка, входящему в один из таких миноров, так что мы можем применить к этому минору второго порядка неравенство (4.22). Но внедиагональный эле-

¹⁾ Это теорема Шура (см. Беллман [1969], стр. 231). В дальнейшем нам потребуется более сильное утверждение, а именно что приведение к треугольному виду можно всегда осуществить таким образом, чтобы собственные значения при этом расположились на диагонали в любом наперед заданном порядке. В самом деле, если $J = S^{-1}AS$ — жорданова каноническая форма матрицы A , то для нее желаемое расположение собственных значений на диагонали можно получить с помощью преобразования P перенумерации осей, приводящего к матрице $P^{-1}JP = (SP)^{-1}ASP$. Применяя теперь процесс ортогонализации Грама — Шмидта к столбцам матрицы SP , получаем унитарную матрицу $U = SPT$, где T — верхняя треугольная матрица. Поэтому матрица $U^*AU = T^{-1}P^{-1}JPT$ имеет требуемый вид.

²⁾ Это следует из треугольности матрицы A : если

$$\left| \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix}^{-1} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right| \leq \alpha \left| \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} \right|$$

для всех v_1, v_2 , то при $v_2 = 0$ получаем $|A_1^{-1}v_1| \leq \alpha |v_1|$.

мент матрицы

$$\begin{bmatrix} \kappa_i - z & A_{i, i+1} \\ 0 & \kappa_{i+1} - z \end{bmatrix}^{-1}$$

равен $-A_{i, i+1}/[(\kappa_i - z)(\kappa_{i+1} - z)]$, поэтому

$$|A_{i, i+1}| \leq C_R \frac{|z - \kappa_i| |z - \kappa_{i+1}|}{|z| - 1}. \quad (4.27)$$

Чтобы установить более полезную оценку

$$|A_{i, i+1}| \leq 16C_R \max\{\gamma_{i, i+1}, |\kappa_i - \kappa_{i+1}|\}, \quad (4.28)$$

где

$$\gamma_{ij} = \min\{1 - |\kappa_i|, 1 - |\kappa_j|\}$$

покажем сначала, что

$$|A_{i, i+1}| \leq 16C_R \max\{1 - |\kappa_{i+1}|, |\kappa_i - \kappa_{i+1}|\}, \quad (4.28a)$$

воспользовавшись для этого свободой в выборе z . Полагая $z = 3$, получаем $|A_{i, i+1}| \leq 8C_R$, так что неравенство (4.28a) выполняется при $|\kappa_{i+1}| \leq 1/2$. Далее, если $|\kappa_{i+1}| > 1/2$, то положим $z = t/\bar{\kappa}_{i+1}$, где $t > 1$. Подставляя это значение z в (4.27), находим

$$|A_{i, i+1}| \leq C_R \frac{(t - |\kappa_{i+1}|^2) |t - \kappa_i \bar{\kappa}_{i+1}|}{|\kappa_{i+1}| (t - |\kappa_{i+1}|)}.$$

При $t \rightarrow 1$ получаем отсюда, что $|A_{i, i+1}| \leq 3C_R |1 - \kappa_i \bar{\kappa}_{i+1}|$. Но

$$\begin{aligned} |1 - \kappa_i \bar{\kappa}_{i+1}| &= |1 - |\kappa_{i+1}|^2 + \bar{\kappa}_{i+1}(\kappa_{i+1} - \kappa_i)| \leq \\ &\leq 3 \max\{1 - |\kappa_{i+1}|, |\kappa_i - \kappa_{i+1}|\}, \end{aligned}$$

так что неравенство (4.28a) справедливо также и в этом случае. Неравенство (4.28) следует из (4.28a) в силу симметрии последнего относительно κ_i и κ_{i+1} .

Мы можем теперь доказать справедливость оценки (4.23) для каждого элемента $B_{i, i+1}$ первой верхней диагонали при $C_B \leq 16C_R$. Если $\gamma_{i, i+1} \geq |\kappa_i - \kappa_{i+1}|$, то оценка (4.23) верна уже для $A_{i, i+1}$ и никакого дополнительного преобразования не требуется; в противном случае элемент $A_{i, i+1}$ можно обратить в нуль с помощью ограниченного преобразования подобия. Действительно, в общем случае элемент A_{ij} при $j > i$ аннулируется преобразованием S_{ij} (т. е. $(S_{ij} A S_{ij}^{-1})_{ij} = 0$), где $S_{ij} = I + T_{ij}$, а T_{ij} — матрица, у которой i, j -й элемент равен $A_{ij}/(\kappa_i - \kappa_j)$, а все остальные элементы равны нулю. Если есть надобность в преобразовании $S_{i, i+1}$, то в силу (4.28) оно и обратное к нему $S_{i, i+1}^{-1} = I - T_{i, i+1}$ будут ограничены по норме одной и той же для всех $A \in \mathcal{F}$ константой. Применяя не более чем $p-1$ таких

преобразований, мы обеспечим выполнение условия (S) для элементов первой верхней диагонали при $C_S \leq 1 + 16C_R$.

Для продолжения этого процесса нам и нужна следующая ключевая лемма:

Если семейство матриц A размера $m \times m$, имеющих верхнюю треугольную форму, удовлетворяет резольвентному условию с константой C_1 и если все их внедиагональные элементы, кроме A_{1m} , удовлетворяют неравенству $|A_{ij}| \leq C_2 \gamma_{ij}$, то

$$|A_{1m}| \leq 16C_1 [1 + (m-2)C_2^2]^{1/2} \max \{\gamma_{1m}, |\kappa_1 - \kappa_m|\}. \quad (4.29)$$

Для случая $m = 2$, когда единственным внедиагональным элементом является A_{12} , доказательство уже имеется. Для общего случая оно получается сведением его к случаю $m = 2$. Для этого переставим в матрице $A - zI$ столбцы и строки с номерами 2 и m и запишем результат E в следующем виде:

$$E = \begin{bmatrix} E_1 & E_2 \\ E_3 & E_4 \end{bmatrix} = \left[\begin{array}{cc|cccc} \zeta_1 & A_{1m} & A_{13} & \dots & A_{1,m-1} & A_{12} \\ 0 & \zeta_m & 0 & \dots & 0 & 0 \\ \hline 0 & A_{3m} & & & & \\ \vdots & \vdots & & & & \\ \vdots & \vdots & & & & \\ 0 & A_{m-1,m} & & & E_4 & \\ 0 & A_{2m} & & & & \end{array} \right]$$

здесь $\zeta_i = \kappa_i - z$; выписывать E_4 в явном виде нет необходимости. Такая перестановка не нарушает условия резольвентности. Очевидно, что $E_3 E_2 = 0$ и потому $E_3 E_1^{-1} E_2 = 0$. Следовательно, если представить E в виде произведения двух треугольных матриц, $E = LU$, то

$$L = \begin{bmatrix} I & 0 \\ E_3 E_1^{-1} & I \end{bmatrix}, \quad U = \begin{bmatrix} E_1 & E_2 \\ 0 & E_4 \end{bmatrix}.$$

Поэтому из резольвентного условия следует, что для любого вектора u

$$|E^{-1}u|^2 = u^* (L^{-1})^* (U^{-1})^* U^{-1} L^{-1} u \leq (C_1/\zeta)^2 |u|^2,$$

где $\zeta = |z| - 1$. Полагая $u = Lv$, где все компоненты вектора v , кроме первых двух, равны нулю (обозначим вектор, соответствующий этим двум компонентам, через v_1), получаем

$$|E_1^{-1} v_1|^2 \leq (C_1/\zeta)^2 |Lv|^2. \text{ Но}$$

$$|Lv|^2 \leq |v_1|^2 \left(1 + \sum_{i=2}^{m-1} A_{im}^2 / \zeta_m^2 \right) \leq [1 + (m-2)C_2^2] |v_1|^2,$$

так что E_1 удовлетворяет неравенству (4.22), в котором C_R заменено на $C_1[1 + (m - 2)C_2^2]^{1/2}$; тем самым общий случай сведен к случаю $m = 2$.

Теперь для проверки выполнения условия (S) осталось сделать совсем немного. Если оценка (4.23) уже установлена для k -й верхней диагонали, то каждый элемент $(k + 1)$ -й диагонали является верхним правым элементом одной из матриц, к которым может быть применена лемма. Далее, используя преобразования подобия $S_{i, i+k+1}$ аналогично тому, как это делалось выше, распространим оценку (4.23) на эту верхнюю диагональ. То, что такой процесс приводит к нужному результату, связано, конечно, с тем обстоятельством, что эти преобразования помимо обращения в нуль наперед заданного элемента изменяют только те элементы, которые лежат на более высоких диагоналях. Отметим также, что в случае каждого применения такого преобразования константа в резольвентном условии умножается на квадрат его нормы.

Этим завершается доказательство теоремы Крайса о матрицах. Ключевая часть доказательства, а именно то, что из (R) следует (S), основана на результатах Мортон и Шехтера [1965]. Дальнейшее обсуждение теоремы и в частности резольвентного условия (R) можно найти у Мортон [1964] и у Миллера и Стрэнга [1965].

§ 4.10. Критерий устойчивости Баченана

Теорема предыдущего параграфа еще не дает непосредственно пригодных для практики критериев устойчивости. Условие (S) обеспечивает лишь существование преобразования подобия, приводящего A к такой форме, которая позволяет судить об устойчивости, и хотя при доказательстве это преобразование определялось конструктивным образом, такой процесс в общем случае невозможно осуществить на практике ввиду его сложности. Следующий наш шаг состоит в том, чтобы научиться решать вопрос об устойчивости семейства матриц, используя только унитарные преобразования, приводящие к треугольной форме. Наилучший результат в этом направлении принадлежит Баченану [1963б], который показал, что если подходящим образом расположить собственные значения на диагонали матрицы, приведенной к треугольной форме, то неравенства типа (4.28) или (4.29) будут как необходимыми, так и достаточными условиями устойчивости.

Назовем последовательность комплексных чисел $\kappa_1, \kappa_2, \dots, \kappa_r$ *кучной* с константой кучности K , если

$$|\kappa_r - \kappa_s| \leq K |\kappa_l - \kappa_m| \quad \text{для} \quad l \leq r \leq s \leq m.$$

Нетрудно показать, что любое множество из p чисел можно упорядочить таким образом, что оно станет кучным с константой $K \leq 2^p$. Действительно, выберем первое число произвольно, вторым назовем ближайшее к первому, третьим — ближайшее из оставшихся чисел к первым двум и т. д. Тогда

$$|\kappa_{i-1} - \kappa_i| \leq |\kappa_{i-1} - \kappa_j| \quad \text{при } j \geq i.$$

Так как

$$|\kappa_r - \kappa_s| \leq |\kappa_r - \kappa_m| + |\kappa_s - \kappa_m|,$$

то при $l \leq r \leq s \leq m$

$$\begin{aligned} |\kappa_r - \kappa_m| &\leq |\kappa_{r-1} - \kappa_r| + |\kappa_{r-1} - \kappa_m| \leq \\ &\leq 2|\kappa_{r-1} - \kappa_m| \leq \dots \leq 2^{r-l} |\kappa_l - \kappa_m|. \end{aligned}$$

Аналогично $|\kappa_s - \kappa_m| \leq 2^{s-l} |\kappa_l - \kappa_m|$, и потому

$$|\kappa_r - \kappa_s| \leq (2^{r-l} + 2^{s-l}) |\kappa_l - \kappa_m| \leq 2^p |\kappa_l - \kappa_m|.$$

Теперь покажем, что справедлива следующая теорема.

Пусть $\mathcal{F}$ — семейство матриц A , имеющих верхнюю треугольную форму, причем собственные значения каждой из них расположены кучно с константой K_1 . Это семейство устойчиво тогда и только тогда, когда для собственных значений κ_i выполняется условие фон Неймана $|\kappa_i| \leq 1$, а для недиагональных элементов справедлива оценка

$$|A_{ij}| \leq K_2 \max\{1 - |\kappa_i|, 1 - |\kappa_j|, |\kappa_i - \kappa_j|\} \quad (4.30)$$

с некоторой константой K_2 , не зависящей от A .

Эта теорема почти сразу следует из теоремы Крайса. Нам нужно лишь показать, что при условии кучности внедиагональные элементы матриц $S_{ij}AS_{ij}^{-1}$ и $S_{ij}^{-1}AS_{ij}$ ($j > i$) удовлетворяют неравенствам (4.30), быть может с другой константой. Действительно, если семейство устойчиво, то в результате применения преобразования S_{ij} получим для преобразованного семейства матриц оценку (4.23) и, следовательно, (4.30), так что условия теоремы необходимы для устойчивости. С другой стороны, мы имеем неравенство

$$1 - |\kappa_i| \leq 1 - |\kappa_j| + |\kappa_i - \kappa_j|,$$

которое симметрично относительно i и j , так что если обозначить через Γ_{ij} максимум в правой части неравенства (4.30), то

$$\Gamma_{ij} \leq 2 \max\{\gamma_{ij} |\kappa_i - \kappa_j|\} \leq 2\Gamma_{ij}.$$

Следовательно, если выполняется (4.30), то $|A_{ij}| \leq 2K_2 \max\{\gamma_{ij}, |\kappa_i - \kappa_j|\}$, и последовательное применение преобразований S_{ij} ,

изменяющих верхние диагонали матрицы A , позволит получить (4.23) и тем самым доказать устойчивость. Поэтому условия теоремы достаточны для устойчивости.

Итак, предположим, что для элементов матрицы A справедлива оценка (4.30). Тогда элементы матрицы $S_{ij}AS_{ij}^{-1}$ отличаются от элементов матрицы A только в i -й строке правее ij -го элемента, где они равны $\hat{A}_{iv} = A_{iv} + t_{ij}A_{jv}$, и в j -м столбце выше этого элемента, где они равны $\hat{A}_{\mu j} = A_{\mu j} - t_{ij}A_{\mu i}$, а $|t_{ij}|$ ограничены постоянной, не зависящей от A . Поскольку $\mu \leq i \leq j \leq v$, то

$$|A_{jv}| \leq K_2 \Gamma_{jv} \leq 2K_2 \max\{\gamma_{jv} |\kappa_j - \kappa_v|\} \leq \\ \leq 2K_2 \max\{1 - |\kappa_v|, K_1 |\kappa_i - \kappa_v|\} \leq 2K_2(1 + K_1) \Gamma_{iv},$$

так что

$$|\hat{A}_{iv}| \leq K_2[1 + 2|t_{ij}(1 + K_1)|] \Gamma_{iv}.$$

Аналогично оцениваются $A_{\mu i}$ и $\hat{A}_{\mu j}$. Следовательно, для элементов матрицы $S_{ij}AS_{ij}^{-1}$ оценка (4.30) снова выполняется, если заменить K_2 коэффициентом при Γ_{iv} из последнего неравенства. Для матриц $S_{ij}^{-1}AS_{ij}$ изменится только знак у t_{ij} . Теорема доказана.

Как уже отмечалось выше, любое множество из p чисел может быть сделано кучным и любая матрица может быть приведена к верхней треугольной форме с заданным расположением собственных значений при помощи некоторого унитарного преобразования, так что в принципе можно считать, что любое семейство матриц задано в требуемом в теореме виде, и исследовать его на устойчивость, проверяя, выполнены ли условие фон Неймана и оценка (4.30). В конце следующего параграфа мы проиллюстрируем этот подход на примере разностной схемы для волнового уравнения.

Может возникнуть сомнение, не являются ли неравенства (4.30) необходимыми или достаточными для устойчивости без каких-либо дополнительных условий. Следующие два контр-примера показывают, что это не так.

Рассмотрим сначала матрицу:

$$\begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix},$$

которая удовлетворяет условию (4.30); n -я степень этой матрицы совпадает с ней в каждом элементе, за исключением верхнего правого, который равен $n - 1$, поэтому степени этой матрицы неограничены,

В качестве второго примера рассмотрим однопараметрическое семейство матриц

$$\begin{bmatrix} \alpha & 1 & -(2\alpha)^{-1} \\ 0 & -\alpha & 1 \\ 0 & 0 & \alpha \end{bmatrix}, \quad 1/2 < \alpha < 1.$$

Верхний правый элемент не удовлетворяет условию (4.30) при $\alpha \rightarrow 1$, однако, как нетрудно проверить, резольвентное условие в данном случае выполняется, так что это семейство устойчиво.

§ 4.11. Дальнейшие примеры достаточных условий

Применение критерия Баченана требует предварительно весьма сложного в общем случае приведения матриц к треугольной форме специального вида, и поэтому мы рассмотрим в этом параграфе ряд более слабых достаточных условий, которые проще применять на практике. Для этой цели мы снова обратимся к матрицам перехода G , не забывая однако о связи между устойчивостью матриц G и устойчивостью семейства матриц $\{e^{-\alpha \Delta t} G\}$, установленной в начале § 4.9.

Сначала сделаем одно замечание общего характера. Предположим, что функция $G(\Delta t, k)$ равномерно по k удовлетворяет условию Липшица при $\Delta t = 0$ в том смысле, что $G(\Delta t, k) = G(0, k) + O(\Delta t)$ при $\Delta t \rightarrow 0$, где $O(\Delta t)$ не зависит от k . Тогда устойчивость разностной схемы можно определить, рассматривая только $G(0, k)$. Это следует из теоремы о малых возмущениях, изложенной в § 3.9. Заметим однако, что это не следует лишь из сравнения $G(0, k)$ и $e^{-\alpha \Delta t} G(\Delta t, k)$, аналогичного тому, которое проводилось выше для $G(\Delta t, k)$, так как $G(\Delta t, k)$ в общем случае не ограничено величиной $e^{\alpha \Delta t} G(0, k)$ ни при каком выборе α .

Если в дополнение к условию Липшица удастся еще показать, что $\|G(0, k)\| \leq 1$, то отсюда сразу следует устойчивость. В этой связи полезно отметить, что норма произвольной матрицы G равняется корню квадратному из спектрального радиуса матрицы G^*G . Действительно, так как

$$\|G\|^2 = \max_{|v|=1} |Gv|^2 = \max_{|v|=1} v^* G^* G v$$

и матрица G^*G нормальна, то, разлагая v по системе ее собственных векторов, получаем, что указанный максимум достигается тогда, когда v является нормированным собственным вектором G^*G , отвечающим ее наибольшему собственному значению.

В гл. 9 даны приложения этих результатов к уравнению переноса и, в частности, к его аппроксимации, получаемой так назы-

ваемым методом сферических гармоник. Например, для одного разностного уравнения, предложенного Фридрихсом, матрица перехода имеет вид

$$G(\Delta t, k) = (I + A \Delta t) \cos k \Delta z + iB \sin k \Delta z,$$

где A и B — фиксированные эрмитовы матрицы, порядок которых равен порядку сферических гармоник, используемых в данном приближении. Из этого примера видно, между прочим, что иногда может оказаться полезной замена переменной k на новую переменную $\xi = (k_1 \Delta x_1, k_2 \Delta x_2, \dots, k_d \Delta x_d)$. В самом деле, если $\Delta z = c \Delta t$, то функция $G(\Delta t, k)$ в этом примере не удовлетворяет условию Липшица при $\Delta t = 0$ равномерно по k , но если рассматривать ее как функцию $\tilde{G}(\Delta t, \xi)$ от Δt и ξ , то она будет обладать этим свойством. Более того, матрица $\tilde{G}(\Delta t, \xi)$ не является нормальной при $\Delta t \neq 0$, поскольку A и B не перестановочны между собой, но матрица $\tilde{G}(0, \xi)$ нормальна, так что условие фон Неймана в этом случае не только необходимо, но и достаточно для устойчивости. Может случиться, что $G(\Delta t, k)$ вообще не удовлетворяет условию Липшица при $\Delta t = 0$, в то время как $G^*(\Delta t, k)G(\Delta t, k)$ удовлетворяет ему. В гл. 9 имеется пример, где $G(\Delta t, k)$ — полином от $\sqrt{\Delta t}$ и тем не менее вопрос об устойчивости $G(\Delta t, k)$ удастся свести к исследованию на устойчивость только матриц $G(0, k)$.

Рассмотрим ряд достаточных условий устойчивости.

Первое достаточное условие. После нормальных матриц G наиболее простыми являются так называемые равномерно диагонализуемые матрицы G ; это означает, что для матрицы G существует такая матрица T , что матрица $\Lambda = T^{-1}GT$ диагональна, а матрицы T и T^{-1} ограничены по норме постоянной, не зависящей от k и от достаточно малых Δt ; таким образом,

$$GT = T \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_p \end{bmatrix} = T\Lambda.$$

Ясно, что условие фон Неймана в этом случае достаточно для устойчивости, так как $G^n = T\Lambda^n T^{-1}$. Чтобы проверить, обладает ли G указанным свойством, мы должны найти все ее собственные векторы; последние образуют столбцы матрицы T , как это видно из равенства $GT = T\Lambda$. Предположим, что они нормированы, так что $\|T\| \leq p$, поскольку норма любой $p \times p$ -матрицы не превосходит p раз взятого наибольшего абсолютного

значения ее элементов¹⁾. Далее, если Δ^2 — определитель матрицы T^*T , т. е. определитель Грама для системы нормированных собственных векторов, то

$$|(T^{-1})_{ij}| \leq \frac{1}{\Delta}.$$

В самом деле, по правилу Крамера $(T^{-1})_{ij}$ равняется отношению алгебраического дополнения элемента T_{ji} к определителю матрицы T , а абсолютное значение любого определителя не превосходит произведения длин векторов, образующих его столбцы (или строки). Из последнего неравенства следует, что $\|T^{-1}\| \leq \leq p/\Delta$, и мы приходим к следующему выводу.

Если матрица G имеет полную систему собственных векторов и существует такая постоянная δ , что $\Delta > \delta > 0$, где Δ^2 — определитель Грама для системы нормированных собственных векторов, то условие фон Неймана не только необходимо, но и достаточно для устойчивости.

Для сравнения с критерием Баченана заметим, что если семейство матриц A имеет верхнюю треугольную форму с кучно расположенными собственными значениями, то для их равномерной диагнолизуемости необходимо и достаточно, чтобы

$$|A_{ij}| \leq C |A_{ii} - A_{jj}|,$$

где C не зависит от A . Достаточность этого условия следует из приведенного при доказательстве теоремы Крайса построения преобразования подобия S с помощью операторов S_{ij} , аннулирующих элементы A_{ij} , а также из того факта, что это условие инвариантно относительно преобразований S_{ij} , если собственные значения расположены кучно. Необходимость доказать несколько сложнее, и мы не будем здесь на этом останавливаться.

Эта теорема может быть применена к явному разностному уравнению для задачи о звуковых волнах. Как мы упоминали в § 4.8, матрица перехода G , имеющая вид (4.20), для этой задачи не является нормальной; но если $v^{(1)}$ и $v^{(2)}$ — нормирован-

¹⁾ Если M — наибольший по модулю элемент $p \times p$ -матрицы A и v — произвольный вектор, то в силу неравенства Шварца

$$\left| \sum_{j=1}^p A_{ij} v_j \right|^2 \leq \left[\sum_{j=1}^p |A_{ij}|^2 \right] \left[\sum_{j=1}^p |v_j|^2 \right] \leq p M^2 |v|^2.$$

Поэтому

$$|Av|^2 \leq p^2 M^2 |v|^2 \quad \text{и} \quad \|A\| \leq pM.$$

Мы могли бы, конечно, использовать здесь более точный результат, что норма $p \times p$ -матрицы ограничена произведением $p^{1/2}$ на максимум норм ее столбцов (или строк); доказательство аналогично.

ные собственные векторы матрицы G (которые легко вычисляются), то

$$|\det(\mathbf{v}^{(1)}, \mathbf{v}^{(2)})| = 1 - \left(\frac{c \Delta t}{\Delta x}\right)^2 \sin^2\left(\frac{k \Delta x}{2}\right)$$

(при условии что выполнено условие фон Неймана $c\Delta t/\Delta x \leq 1$). Модуль этого определителя ограничен снизу положительной константой, если

$$\frac{c \Delta t}{\Delta x} = \text{const} < 1 \quad (4.31)$$

при $\Delta t \rightarrow 0$ и $\Delta x \rightarrow 0$. Таким образом, мы приходим к хорошо известному выводу о том, что если выполнено неравенство (4.31), то рассматриваемые разностные уравнения устойчивы.

Второе достаточное условие. Часто найти собственные векторы матрицы G , что требуется в предыдущем условии, — дело весьма затруднительное. На практике проще всего применять такие условия, которые требуют лишь знания собственных значений. Простейшим из них является следующее:

Если элементы матриц $G(\Delta t, \mathbf{k})$ равномерно ограничены при всех $0 < \Delta t < \tau$ и всех $\mathbf{k} \in \mathcal{L}$ и все их собственные значения λ_i , за исключением быть может одного, принадлежат некоторому кругу, лежащему внутри единичного круга, т. е.

$$0 < \Delta t < \tau$$

$$|\lambda_i| \leq \gamma < 1 \quad \text{при} \quad \mathbf{k} \in \mathcal{L}$$

$$i = 2, \dots, p,$$

то условие фон Неймана не только необходимо, но и достаточно для устойчивости.

Это немедленно следует из критерия Баченана. В самом деле, если матрица G приведена к верхней треугольной форме с кучным расположением собственных значений, то все ее внедиагональные элементы не превосходят по модулю некоторого M . Далее, не нарушая общности, можно считать исключительное собственное значение первым. Поэтому неравенства (4.30) удовлетворяются с $K_2 \leq M/(1 - \gamma)$, и условие фон Неймана, сводящееся к условию $|\lambda_1| \leq 1 + O(\Delta t)$ на это собственное значение, оказывается достаточным для устойчивости.

Применим для примера эту теорему к пятиточечному разностному уравнению

$$\frac{3}{2} \frac{u_j^{n+1} - u_j^n}{\Delta t} - \frac{1}{2} \frac{u_j^n - u_j^{n-1}}{\Delta t} = \sigma \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{(\Delta x)^2} \quad (4.32)$$

(схема 9 из таблицы гл. 8), которое применяется для решения одномерного уравнения теплопроводности. Это трехслойная

формула, но ее можно свести к двухслойной системе¹⁾ при помощи подстановки $u_j^{n-1} = v_j^n$. Эта система имеет вид

$$\frac{3}{2} \frac{u_j^{n+1} - u_j^n}{\Delta t} - \frac{1}{2} \frac{u_j^n - v_j^n}{\Delta t} = \sigma \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{(\Delta x)^2}, \quad (4.33)$$

$$v_j^{n+1} = u_j^n.$$

Ее матрица перехода равна

$$G(\Delta t, k) = \begin{bmatrix} \frac{4}{3 + 8\alpha \sin^2 \beta} & \frac{-1}{3 + 8\alpha \sin^2 \beta} \\ 1 & 0 \end{bmatrix}, \quad (4.34)$$

где для краткости положено

$$\alpha = \frac{\sigma \Delta t}{(\Delta x)^2}, \quad \beta = \frac{1}{2} k \Delta x.$$

Собственные значения этой матрицы суть

$$\lambda_1, \lambda_2 = \frac{2 \pm \sqrt{1 - 8\alpha \sin^2 \beta}}{3 + 8\alpha \sin^2 \beta},$$

и легко видеть, что условие фон Неймана выполняется для произвольных значений α . На рис. 4.2 помещены графики, показы-

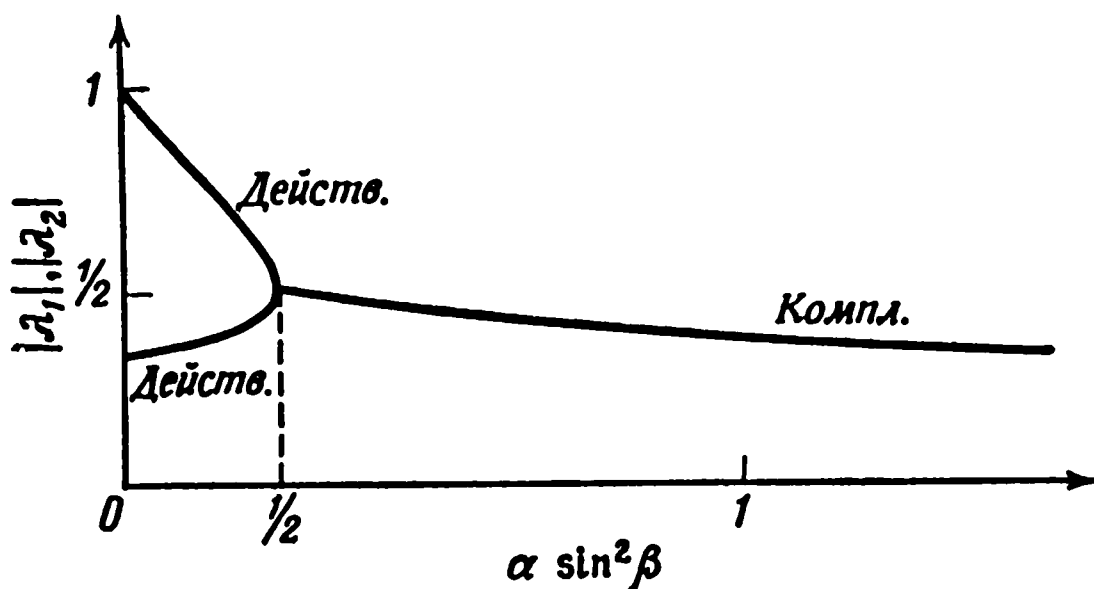


Рис. 4.2. Собственные значения матрицы перехода системы разностных уравнений (4.33).

вающие зависимость $|\lambda_1|$ и $|\lambda_2|$ от $\alpha \sin^2 \beta$. Собственные векторы матрицы G равны

$$\varphi^{(1)}, \varphi^{(2)} = A \left(\frac{1}{2 \mp \sqrt{1 - 8\alpha \sin^2 \beta}} \right),$$

¹⁾ Мы еще вернемся к этому примеру в начале гл. 7, где исследуются многослойные формулы общего вида. Как мы увидим если уравнения сведутся к эквивалентной двухслойной системе, то исследование устойчивости проводится точно так же, как в настоящей главе.

где A — нормирующий множитель. Определитель матрицы нормированных собственных векторов равен

$$\Delta = 2A^2 \sqrt{1 - 8\alpha \sin^2 \beta}.$$

Если $\alpha > 1/8$, то Δ может обращаться в нуль при некоторых значениях β ; следовательно, мы не можем воспользоваться первым достаточным условием устойчивости. Однако второе условие применимо, так как $|\lambda_2| \leq 1/2$ для всех положительных α и всех β (это видно из рис. 4.2). Таким образом, система (4.33) всегда устойчива.

Достаточное условие Като. Оба предыдущих условия являются частными случаями следующего условия, принадлежащего Като [1960].

Предположим, что нормы матриц $G(\Delta t, k)$ равномерно ограничены и выполнено условие фон Неймана. Предположим также, что для каждой матрицы G можно построить внутри круга $|\lambda| \leq R$, где R — спектральный радиус G , такую замкнутую спрямляемую кривую Γ , что ее длина равномерно ограничена, а ее расстояние до спектра G не меньше некоторой положительной константы, одной и той же для всех G . Тогда разностная схема будет устойчивой, если существует такая не зависящая от Δt и k постоянная $\delta > 0$, что каждое из различных собственных значений λ_i , $i = 1, 2, \dots, q$, лежащих вне Γ , имеет индекс¹⁾ 1 и либо (1) расстояние от λ_i до всех других собственных значений больше δ , (2) либо для множества собственных значений λ_i , не удовлетворяющих условию (1), полная система соответствующих нормированных собственных векторов имеет определитель Грама Δ^2 , для которого $\Delta > \delta$.

В большинстве практически важных случаев Γ будет просто окружностью. Этого всегда достаточно, если рассматривается отдельная матрица; в то же время при исследовании семейства матриц можно получить более тонкие результаты, используя кривые более общего вида.

Основные моменты доказательства ясны. Приводим каждую матрицу G к верхней треугольной форме с кучным расположением собственных значений, причем по предположению можно расположить сначала те λ_i , которые удовлетворяют условию (2).

¹⁾ Если λ — собственное число матрицы A , то множество векторов v , для которых $(A - \lambda)^n v = 0$ при некотором натуральном n , образует так называемое *инвариантное подпространство*, соответствующее собственному значению λ . Наименьшее значение n , при котором это уравнение удовлетворяется для всех векторов инвариантного подпространства, называется *индексом* собственного значения λ . Если индекс равен единице, то в инвариантном подпространстве можно выбрать базис из собственных векторов.

Тогда, как нетрудно видеть, все элементы, находящиеся выше прочих собственных значений, удовлетворяют неравенству (4.30). Оставшийся квадратный минор можно диагонализировать с помощью равномерно ограниченных преобразований подобия аналогично тому, как это делалось при выводе первого достаточного условия.

Часто оказывается полезным следующее достаточное условие, предложенное Лаксом и Вендроффом [1964].

Если матрицы $G(\Delta t, k)$ таковы, что для любого вектора v

$$|v^* G v| \leq [1 + O(\Delta t)] |v|^2,$$

то соответствующая разностная схема устойчива.

Доказательство состоит просто в применении теоремы Крайса в форме резольвентного условия (4.22). Выбирая α так, чтобы для $A = \exp(-\alpha \Delta t) G$ выполнялось условие $|v^* A v| \leq |v|^2$, мы видим, что все собственные значения A лежат в замкнутом единичном круге. Следовательно, $(A - zI)^{-1}$ существует при $|z| > 1$, и если w — произвольный вектор и $v = (A - zI)^{-1} w$, то

$$|v| |w| \geq |v^* w| = |v^* (A - zI) v| \geq (|z| - 1) |v|^2,$$

т. е.

$$|v| = |(A - zI)^{-1} w| \leq |w| (|z| - 1)^{-1}.$$

Поэтому семейство матриц A удовлетворяет резольвентному условию и устойчивость следует из теоремы Крайса о матрицах.

Положим $\rho(A) = \max \{|v^* A v| \mid |v| = 1\}$. Ясно, что $\rho(A) \leq \|A\|$. Равенство здесь имеет место для нормальных матриц A . Поэтому, разлагая в общем случае матрицу A на эрмитову и косоэрмитову, мы видим, что $\|A\| \leq 2\rho(A)$. Далее, можно показать (см. Пирси [1966]), что если $\rho(A) \leq 1$, то $\rho(A^n) \leq 1$ для любого натурального n . Таким образом, если удовлетворяется условие Лакса — Вендроффа, то соответствующее семейство матриц A не просто ограничено, а ограничено числом 2.

Если ни одно из достаточных условий типа приведенных выше не работает, то можно попытаться использовать критерий Баченана. Рассмотрим в качестве примера предельный случай $c\Delta t = \Delta x$ для явной разностной схемы, аппроксимирующей волновое уравнение. Поскольку матрица G в этом случае дается формулой (4.20), то ее собственные значения равны

$$\lambda_1, \lambda_2 = 1 - \frac{1}{2} a^2 \pm \frac{1}{2} i a \sqrt{4 - a^2},$$

где $a = (2c\Delta t/\Delta x)\sin(k\Delta x/2)$; им соответствуют нормированные собственные векторы

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1/2(\sqrt{4-a^2} + ia) \end{bmatrix} \quad \text{и} \quad \frac{1}{\sqrt{2}} \begin{bmatrix} -1/2(\sqrt{4-a^2} + ia) \\ 1 \end{bmatrix}.$$

Матрица G приводится к верхней треугольной форме U^*GU с помощью унитарного преобразования

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1/2(\sqrt{4-a^2} - ia) \\ 1/2(\sqrt{4-a^2} + ia) & 1 \end{bmatrix},$$

откуда

$$U^*GU = \begin{bmatrix} 1 - 1/2a^2 + 1/2ia\sqrt{4-a^2} & 1/2a^2(ia - \sqrt{4-a^2}) \\ 0 & 1 - 1/2a^2 - 1/2ia\sqrt{4-a^2} \end{bmatrix}.$$

Абсолютная величина собственных значений равна 1, а расстояние между ними $a\sqrt{4-a^2}$ стремится к нулю, когда $a^2 \rightarrow 4$. Но абсолютное значение недиагонального элемента при этом равно 4, так что система в этом случае оказывается неустойчивой.

Стоит отметить, что Курант, Фридрихс и Леви [1928] доказали, что в этом случае имеет место устойчивость, однако при этом они допускали только гладкие начальные данные. В гл. 10 приводится пример неограниченного решения разностных уравнений (для доказательства неустойчивости обычно бывает проще построить такой пример, чем применять критерий Баченана); начальные значения в этом примере сильно осциллируют при $\Delta x \rightarrow 0$.

На этом мы закончим изучение разностных схем для общих дифференциальных уравнений с постоянными коэффициентами и периодическими граничными условиями. Изложенная здесь теория является основой для всего дальнейшего, и мы будем часто на нее ссылаться. В частности, на теореме Крайса о матрицах основано обсуждение в следующей главе задач с переменными коэффициентами, в то время как условие фон Неймана окажется существенным с практической точки зрения при исследовании разностных схем во второй части книги. В случае когда известно, что разностная схема условно устойчива (т. е. устойчива лишь при некоторых значениях $\Delta t/\Delta x$ или при каких-нибудь аналогичных условиях), то условие фон Неймана почти всегда правильно описывает область устойчивости, и специальное исследование обычно бывает необходимо только на границе этой области, как например в случае явной схемы для волнового уравнения. Имеется однако ряд случаев, когда наше определение устойчивости оказывается недостаточным с практической точки зрения. Один такой случай встретится нам в § 10.4 при изучении совместного распространения звука и тепла, другие приведены в § 9.7, 11.6 и 12.16.

Глава 5

ЛИНЕЙНЫЕ ЗАДАЧИ С ПЕРЕМЕННЫМИ КОЭФФИЦИЕНТАМИ. НЕЛИНЕЙНЫЕ ЗАДАЧИ

§ 5.1. Введение

В предыдущей главе мы подробно рассмотрели вопрос об устойчивости конечноразностных аппроксимаций для линейных дифференциальных уравнений с постоянными коэффициентами. Но в практических задачах коэффициенты редко бывают постоянными, а уравнения очень часто бывают нелинейными. В таких случаях фактически проверяется «локальная» устойчивость линеаризованных уравнений, полученных из исходных нелинейных уравнений, при этом постоянно используется теория для случая постоянных коэффициентов. Вот как это делается. Из нелинейных разностных уравнений относительно функции $u^n(x)$, аппроксимирующей решение $u(x, n\Delta t)$ дифференциальной задачи, получают уравнения первой вариации, производя подстановку $u^n(x) = u(x, n\Delta t) + v^n(x)$, где $v^n(x)$ рассматривается как величина первого порядка малости, и отбрасывая члены второго и более высоких порядков малости. Получающиеся для $v^n(x)$ уравнения линейны, и их коэффициенты зависят от функции $u(x, t)$, которая является величиной нулевого порядка. Затем в каждой точке (x, t) рассматривают систему уравнений, которая получается, если положить коэффициенты постоянными, а именно равными значениями переменных коэффициентов, вычисленным в этой точке, и исследуют ее на устойчивость методами теории для случая постоянных коэффициентов. Успех такого подхода определяется тем, насколько локальная устойчивость в каждой точке пространства переменных x, t связана со сходимостью $u^n(x)$ к решению $u(x, t)$. Практика показывает, что неустойчивость обычно начинается как локальное явление; это наводит на мысль, что указанная связь очень тесная; предположение о такой связи было одной из причин, побудивших фон Неймана разрабатывать теорию устойчивости с использованием преобразования Фурье.

Ясно, что в первую очередь следует рассмотреть линейные уравнения с переменными коэффициентами, и в настоящей главе мы уделим этому вопросу главное внимание. Кроме того мы ограничимся рассмотрением только таких случаев, когда коэффициенты зависят лишь от пространственной переменной x . Основная наша цель состоит в том, чтобы показать, что для не-

скольких важных классов уравнений незначительное усиление условий локальной устойчивости оказывается достаточным для обеспечения устойчивости в целом. Однако эти локальные условия не всегда являются даже необходимыми, и стоит вначале потратить немного времени, чтобы разобраться в этом вопросе.

Прежде всего рассмотрим взаимосвязь между корректностью постановки задач для соответствующих дифференциальных уравнений. Крайс [1962] дал пример уравнения, для которого задача с переменными коэффициентами поставлена корректно, хотя все соответствующие задачи с постоянными коэффициентами поставлены некорректно¹⁾, с другой стороны, он построил пример, в котором все задачи с постоянными коэффициентами поставлены корректно, в то время как задача с переменными коэффициентами не является таковой. Как мы уже отмечали в § 3.9, для некорректно поставленной задачи не существует разностной схемы, которая одновременно была бы согласованной и устойчивой. Поэтому примеры Крайса показывают, что для задач с переменными коэффициентами локальная устойчивость в общем случае не является ни необходимой, ни достаточной для устойчивости в целом. Однако, как выяснил Стрэнг [1966], если для уравнения $\partial u / \partial t = Au$ задача поставлена корректно, то после замены оператора A его главной частью (т. е. суммой членов, содержащих производные наивысшего порядка) для получившегося уравнения все задачи с постоянными коэффициентами будут также корректно поставленными. Поэтому, в частности, задачи для систем первого порядка должны быть локально корректно поставленными.

Стрэнг доказал также аналогичный результат для явных разностных схем. Предположим, что дифференциальное уравнение $\partial u / \partial t = Au$, где A — дифференциальный оператор m -го порядка, аппроксимируется явной схемой $u^{n+1} = C(\Delta t)u^n$. Тогда наивысшая степень отношения $1/\Delta x$, встречающаяся в $C(\Delta t)$ и соответствующая главной части оператора A , равна m . Величину $(1/\Delta x)^m$ надо умножить на Δt , и если выполняется условие $\Delta t = \text{const}(\Delta x)^m$, которое обычно бывает необходимо для устойчивости, то при $\Delta t \rightarrow 0$ $C(\Delta t)$ будет иметь предел $C(0)$,

¹⁾ Следующий аналогичный, но несколько более простой пример принадлежит Стрэнгу [1966]:

$$\frac{\partial u}{\partial t} = Au, \quad \text{где} \quad A = i \frac{\partial}{\partial x} \sin x \frac{\partial}{\partial x} = i \sin x \frac{\partial^2}{\partial x^2} + i \cos x \frac{\partial}{\partial x}.$$

При $x = 0$ получаем $\partial u / \partial t = i \partial u / \partial x$; ясно, что задача для этого уравнения поставлена некорректно. С другой стороны, для любого решения u основной задачи с помощью интегрирования по частям находим $\frac{\partial}{\partial t} (u, u) = (u, Au) + (Au, u) = 0$, так что основная задача поставлена корректно.

соответствующий аппроксимации одной главной части оператора A ; поэтому $C(0)$ называется главной частью оператора $C(\Delta t)$. Справедлива следующая теорема.

Предположим, что явная разностная схема

$$u^{n+1}(x) = (C(\Delta t) u^n)(x) \equiv \sum_{\beta \in N_0} C^\beta(x, \Delta t) u^n(x + \beta \Delta x) \quad (5.1)$$

имеет одинаковое по всем пространственным координатам приращение Δx , которое связано с Δt соотношением типа (4.9). Предположим также, что она устойчива и что при $\Delta t \rightarrow 0$ каждый коэффициент $C^\beta(x, \Delta t)$ сходится к ограниченной функции $C^\beta(x, 0)$. Тогда для каждой внутренней точки a , принадлежащей множеству, на котором эта сходимость равномерна и $C^\beta(x, 0)$ непрерывны, разностная схема

$$U^{n+1}(x) = (C_a(0) U^n)(x) \equiv \sum_{\beta \in N_0} C^\beta(a, 0) U^n(x + \beta \Delta x) \quad (5.2)$$

является устойчивой.

Докажем эту теорему. Предположим, что она неверна в некоторой точке a . Пусть непрерывность и равномерная сходимость имеют место в шаре S_0 с центром в точке a . Тогда для любого, сколь угодно большого, M существуют такие вектор V , гармоника k , приращение Δx и натуральное число N , что ¹⁾ для $x = a$ и $N\Delta t < 1$

$$|\Gamma(x)^N V| > M|V|, \quad (5.3)$$

где

$$\Gamma(x) = \sum_{\beta \in N_0} c^\beta(x, 0) e^{ik \cdot \beta \Delta x}.$$

По непрерывности это неравенство справедливо для всех x из некоторого шара S_1 с центром в точке a , радиус которого мы будем считать строго меньшим, чем радиус шара S_0 ; если мы зафиксируем $\xi = k\Delta x$, то неравенство останется справедливым при $\Delta t \rightarrow 0$ и $\Delta x \rightarrow 0$ и фиксированном N . Пусть теперь $\rho(x)$ — гладкая функция, равная нулю вне S_1 и единице в точке a , и пусть $v(x) = V\rho(x) e^{ik \cdot x}$. Тогда

$$\begin{aligned} (Cv)(x) &= \sum_{\beta \in N_0} c^\beta(x, \Delta t) V\rho(x + \beta \Delta x) e^{ik \cdot (x + \beta \Delta x)} = \\ &= \Gamma(x) v(x) + \varepsilon_1(x, \Delta t), \end{aligned}$$

где ε_1 при достаточно малых Δt обращается в нуль вне некоторого множества, целиком лежащего внутри S_0 , и равномерно

¹⁾ Как и в предыдущей главе, в дальнейшем мы будем, если специально не оговорено иное, использовать символ $|\cdot|$ для обозначения евклидовой нормы вектора, а символ $\|\cdot\|$ для обозначения нормы в пространстве L_2 .

стремится к нулю на этом множестве при $\Delta t \rightarrow 0$. Последовательно применяя оператор C , мы получим

$$(C^N v)(x) = \Gamma(x)^N v(x) + \varepsilon_N(x, \Delta t),$$

где ε_N обладает теми же свойствами, что и ε_1 , поскольку N фиксировано независимо от Δt . Используя далее (5.3) и выбирая Δt достаточно малым, находим, что $\|C^N\| \geq M$, а значит, схема (5.1) неустойчива вопреки нашему предположению.

Используя соображения, аналогичные приведенным, мы рассмотрим в этой главе два класса уравнений, для которых ситуация относительно ясна, а именно, симметричные гиперболические системы и параболические уравнения. Ими охватывается значительная часть важных для физики краевых задач; как хорошо известно, в обоих случаях корректность постановки линейных задач с переменными коэффициентами обеспечивается теоремами чисто локального характера. Кроме того, мы ограничимся только явными разностными схемами, хотя нетрудно понять, как полученные ниже результаты можно распространить на неявные схемы. Как мы увидим ниже в этой главе и в гл. 8, в критерии устойчивости для параболических уравнений по существу входит лишь главная часть оператора. Следовательно, теорема Стрэнга показывает, что локальная устойчивость является необходимой в обоих рассматриваемых случаях.

Эти необходимые критерии устойчивости нужно лишь незначительно усилить, чтобы получить условия, достаточные для устойчивости задачи с переменными коэффициентами. Непосредственно видно, что если $a(x)$ изменяется, то уже первая разность $\Delta[a(x)u(x)]$ содержит дополнительный член $(\Delta a)u$, который может вызвать неустойчивость, если $\Delta x \neq O(\Delta t)$ и $a(x)$ не удовлетворяет условию Липшица, так что дополнительный член оценивается величиной $O(\Delta t)\|u\|$. Для разностей более высокого порядка, например для $\Delta(a\Delta u)$, получается дополнительный член $(\Delta a)(\Delta u)$, который может привести к неприятностям, даже при $\Delta x = O(\Delta t)$. Подобные члены вызывают перемешивание компонент Фурье функции u , аналогичное описанному для случая нелинейных задач у Н. Филлипса [1959], хотя и несколько менее ярко выраженное; хороший пример такого рода имеется у Робертса и Вайса [1966]. Последовательный рост высокочастотных компонент можно устранить введением в разностные схемы некоторой «диссипации». При этом получается схема, у которой собственные значения λ_v матрицы перехода удовлетворяют неравенству вида

$$|\lambda_v| \leq 1 - \delta |k \Delta x|^{2r} \quad \text{при} \quad |k \Delta x| \leq \pi \quad (5.4)$$

для некоторого $\delta > 0$ и некоторого натурального r . Такое условие впервые ввел Фриц Джон [1952] для одного параболического уравнения, где это условие вполне естественно возникает в явных схемах (см. § 5.3); он показал, что оно достаточно для устойчивости в случае достаточно гладких коэффициентов. Затем Лакс [1961] и Лакс и Вендрофф [1962] нашли для симметричных гиперболических систем достаточное условие устойчивости, в котором существенную роль играют неравенства вида (5.4) для собственных значений матрицы перехода. В недавней статье Крайса [1964] оно было сведено к одному лишь условию на собственные значения; этот результат образует ядро § 5.4.

Но прежде чем приступить ко всему этому, посмотрим сначала, как еще можно было бы определить устойчивость. Это особенно уместно сделать здесь, так как именно при рассмотрении задач с переменными коэффициентами различия в определениях, которых придерживаются разные авторы, становятся существенными.

§ 5.2. Другие определения устойчивости

Если оставить в стороне различия в используемых нормах, существуют два широко распространенных определения устойчивости, отличных от того, которое использовалось в нашем изложении, — одно более слабое, а другое более сильное. В первом из них требуется, чтобы разрешающий оператор $C(\Delta t)$ устойчивой разностной схемы удовлетворял неравенству

$$\|C(\Delta t)^n\| \leq \text{const} (\Delta t)^{-\alpha} \leq \text{const} n^{\alpha} \quad \text{при} \quad \begin{cases} 0 < \Delta t < \tau, \\ 0 \leq n \Delta t \leq T \end{cases} \quad (5.5)$$

для некоторого фиксированного конечного $\alpha \geq 0$, в то время как для устойчивости в нашем смысле требуется, чтобы приведенное неравенство было выполнено для $\alpha = 0$. Это определение согласуется с тем экспериментальным фактом, что неустойчивость обычно отличается от устойчивости экспоненциальным, а не полиномиальным ростом ошибок; его использовали многие авторы, в том числе Стрэнг [1960], Вазов и Форсайт [1960], Рябенский и Филиппов [1956].

По сути дела это понятие «слабой» устойчивости порождается более общим определением корректности. Действительно, предполагая, что $u(x, t)$ имеет непрерывные производные по x до порядка p включительно, определим норму $\|u\|_p$ равенством

$$\|u\|_p^2 = \sum_{|v| \leq p} \left\| \partial_{x_1}^{v_1} \partial_{x_2}^{v_2} \dots \partial_{x_d}^{v_d} u \right\|^2, \quad (5.6)$$

где $|\mathbf{v}| = v_1 + v_2 + \dots + v_d$ обозначает порядок производной, среднеквадратичная норма которой стоит под знаком суммы. Назовем задачу *корректно поставленной*, если ее решение и начальный элемент удовлетворяют следующему условию:

$$\|\mathbf{u}(\mathbf{x}, t)\| \leq \text{const} \|\mathbf{u}(\mathbf{x}, 0)\|_p, \quad 0 \leq t \leq T, \quad (5.7)$$

для некоторого фиксированного конечного p и для плотного множества начальных элементов; это соответствует понятию непрерывной зависимости порядка p от начальных данных, введенному Адамаром [1923]. Исходя из этих определений корректности и слабой устойчивости, можно построить аналогичную изложенной в гл. 3 теорию, снова приводящую к теореме об эквивалентности сходимости и устойчивости. Доказательства несколько усложняются, поскольку используются другие нормы; интересующегося подробностями читателя мы отсылаем к упомянутым выше работам. Связь между α и p , фигурирующими в неравенствах (5.5) и (5.7) соответственно, легко установить следующим образом. Из соображений, основанных на принципе равномерной ограниченности, ясно, что для сходимости необходимо, чтобы

$$\|C(\Delta t)^n \mathbf{u}(\mathbf{x}, 0)\| \leq \text{const} \|\mathbf{u}(\mathbf{x}, 0)\|_p$$

для любого элемента $\mathbf{u}(\mathbf{x}, 0)$, имеющего p непрерывных производных, при условии что задача корректно поставлена в указанном выше смысле. Но левая часть этого неравенства зависит только от значений $\mathbf{u}(\mathbf{x}, 0)$ в точках сетки. Для любого $\mathbf{u}(\mathbf{x}, 0)$ мы всегда можем построить имеющий в этих точках те же значения гладкий начальный элемент $\tilde{\mathbf{u}}(\mathbf{x}, 0)$, к которому мы уже можем применить указанное выше неравенство. Далее, легко показать, что $\|\tilde{\mathbf{u}}(\mathbf{x}, 0)\|_p \leq \text{const} (\Delta t)^{-q} \|\mathbf{u}(\mathbf{x}, 0)\|$, где q зависит от p и от размерности d . Следовательно, должно выполняться условие устойчивости вида (5.5).

Если применить эту теорию к задачам с постоянными коэффициентами, рассмотренным в предыдущей главе, то можно получить замечательный по простоте результат, принадлежащий Крайсу [1962], а именно: *если матрицы перехода G равномерно ограничены, то условие фон Неймана (4.18) необходимо и достаточно для слабой устойчивости*. Достаточность легко доказать. Преобразуя G с помощью унитарной матрицы U к верхней треугольной форме, мы получим $UGU^* = D + S$, где

$$D = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_p \end{bmatrix}, \quad S = \begin{bmatrix} 0 & S_{12} & \dots & S_{1p} \\ & \ddots & & \vdots \\ & & \ddots & \vdots \\ 0 & & & S_{p-1,p} \\ & & & 0 \end{bmatrix}.$$

Следовательно,

$$\begin{aligned} \|G^n\| &= \|(UGU^*)^n\| = \|(D+S)^n\| = \\ &= \left\| D^n + \sum_{j=1}^n D^{n-j} S D^{j-1} + \sum_{j=1}^{n-1} \sum_{k=1}^{n-j} D^{n-j-k} S D^{j-1} S D^{k-1} + \dots \right\|. \end{aligned}$$

Но произведение SD имеет тот же вид, что и S , а произведение p матриц вида S обращается в нуль. Поэтому в правой части этого равенства содержится лишь конечное число ненулевых сумм. В то же время $\|S\|$ ограничены, и поскольку $|\lambda_j| \leq 1 + \text{const } \Delta t$, то и $\|D\|^n$ ограничены при $n\Delta t \leq T$. Таким образом, для $n \geq p-1$ имеем

$$\|C^n\| \leq \sum_{j=0}^{p-1} \binom{n}{j} \|D\|^{n-j} \|S\|^j \leq \text{const } n^{p-1}. \quad (5.8)$$

Доказать необходимость условия фон Неймана значительно труднее, и мы отсылаем читателя к статье Крайса.

Применяя это простое алгебраическое условие устойчивости, нетрудно выяснить, как связаны между собой локальная устойчивость и устойчивость в целом. Результат получается неутешительный. В отличие от того, что было доказано для нашего определения устойчивости в § 3.9, добавление оператора порядка $O(\Delta t)$ может сделать слабо устойчивый оператор неустойчивым. Вот простой пример. Предположим, что разностная схема

$$u^{n+1} = (I + S\Delta_+^2) u^n$$

используется для аппроксимации уравнения $\partial u / \partial t = 0$, где вектор u имеет две компоненты, а S — постоянная матрица, у которой только один верхний правый элемент отличен от нуля, а именно равен единице. Тогда собственные значения соответствующей матрицы перехода $G = I + (e^{ik\Delta x} - 1)^2 S$ равны единице, так что схема слабо устойчива. Но если положить в S левый нижний элемент равным Δt , то, как нетрудно проверить, собственные значения соответствующей матрицы перехода будут равны

$$\lambda = 1 \pm (\Delta t)^{1/2} (e^{ik\Delta x} - 1)^2.$$

Следовательно, схема не будет больше устойчивой.

То, как возникают подобные возмущения в задаче с переменными коэффициентами, видно на следующем примере, принадлежащем Крайсу [1962]. Возьмем дифференциальное уравнение $\partial u / \partial t = \partial u / \partial x$; пусть у вектора u две компоненты, $\Delta x = \Delta t$ и разностная аппроксимация имеет вид

$$u^{n+1}(x) = [I + \Delta_+ + V(x) S V^{-1}(x) \Delta_+^2] u^n(x),$$

где Δ_+ — оператор разности вперед, а

$$V(x) = \begin{bmatrix} \cos x & \sin x \\ -\sin x & \cos x \end{bmatrix}, \quad S = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

В случае когда функция $V(x)$ заменяется ее значением $V(x_0)$ в некоторой фиксированной точке x_0 , матрица перехода нашей схемы равна

$$V(x_0)(e^{ik\Delta x}I + (e^{ik\Delta x} - 1)^2 S)V^{-1}(x_0)$$

и очевидным образом удовлетворяет условию фон Неймана при любом x_0 . Рассмотрим теперь случай переменных коэффициентов. Для этого введем зависимое переменное $v(x, t) = V^{-1}(x)u(x, t)$, для которого

$$v^{n+1}(x) = v^n(x) + [V^{-1}(x)\Delta_+ + SV^{-1}(x)\Delta_+^2]V(x)v^n(x).$$

Но

$$\Delta_+[V(x)v(x)] = V(x + \Delta x)\Delta_+v(x) + [\Delta_+V(x)]v(x)$$

и

$$V^{-1}(x)V(x + v\Delta x) = \begin{bmatrix} \cos v\Delta x & \sin v\Delta x \\ -\sin v\Delta x & \cos v\Delta x \end{bmatrix}.$$

Поэтому, выделяя члены порядка $(\Delta x)^2$ и выше, мы получим для v уравнение с постоянными коэффициентами вида

$$v^{n+1}(x) = (A + (\Delta x)^2 B)v^n(x),$$

где оператор B равномерно ограничен и

$$A = \begin{bmatrix} 1 & \Delta x \\ -\Delta x & 1 \end{bmatrix}(I + \Delta_+) + S \begin{bmatrix} 1 & 2\Delta x \\ -2\Delta x & 1 \end{bmatrix}\Delta_+^2 + \\ + 2S \begin{bmatrix} 0 & \Delta x \\ -\Delta x & 0 \end{bmatrix}\Delta_+ + S \begin{bmatrix} 1 & \Delta x \\ -\Delta x & 1 \end{bmatrix}^1).$$

Обозначая преобразование Фурье матричной функции B через $\hat{B}$, получим матрицу перехода

$$G = e^{ik\Delta x} \begin{bmatrix} 1 - 2\alpha\Delta x - 2\beta\Delta x - 2\gamma\Delta x & \Delta x + \alpha + 2\gamma \\ -\Delta x & 1 \end{bmatrix} + (\Delta x)^2 \hat{B},$$

где

$$\alpha = -4\sin^2\left(\frac{1}{2}k\Delta x\right), \quad \beta = 2ie^{-\frac{1}{2}ik\Delta x}\sin\frac{1}{2}k\Delta x, \quad \gamma = \frac{1}{2}e^{-ik\Delta x}.$$

Характеристическое уравнение матрицы из первого слагаемого в выражении для G записывается так:

$$(1 - \mu)(1 - \mu - 2(\alpha + \beta + \gamma)\Delta x) + \Delta x(\Delta x + \alpha + 2\gamma) = 0.$$

¹⁾ Здесь и далее последний член был упущен авторами. — Прим. перев.

Оно дает такие собственные значения:

$$\mu = 1 - (\alpha + \beta + \gamma) \Delta x \pm [-(\alpha + 2\gamma) \Delta x + O[(\Delta x)^2]]^{1/2}.$$

Но собственные значения матрицы G равны $\lambda = e^{ik \Delta x} \mu + O[(\Delta x)^2]$. Вспоминая, что $\Delta t = \Delta x$, мы видим, что условие фон Неймана не удовлетворяется, и потому разностные уравнения для $v(x)$ не являются даже слабо устойчивыми.

Вывод из этих примеров тот, что хотя слабая устойчивость и может быть полезна для задач с постоянными коэффициентами или для случая, когда u является скалярной переменной, она не является достаточно эффективным критерием для систем разностных уравнений с переменными коэффициентами. Даже наше определение устойчивости не всегда оказывается удовлетворительным, когда оно применяется локально, как это было показано Видлундом [1966] на примере, сходном с предыдущим, но несколько более сложном.

Определим теперь «сильную» устойчивость.

Разностная схема называется *сильно устойчивой*, если выполнены следующие условия: (1) при каждом фиксированном Δt оператор $H(\Delta t)$ определен всюду и таков, что если (u, Hu) обозначить через $\|u\|_H^2$, то $K_1^{-1} \|u\|^2 \leq \|u\|_H^2 \leq K_1 \|u\|^2$, (2) решение разностной схемы удовлетворяет условию

$$\|u^{n+1}\|_H \leq (1 + K_2 \Delta t) \|u^n\|_H, \quad (5.9)$$

где K_1 и K_2 — некоторые фиксированные положительные постоянные.

Как мы видели в § 4.9, это определение устойчивости эквивалентно нашему в случае постоянных коэффициентов. Нетрудно убедиться, что в общем случае оно по меньшей мере не слабее, так как если схема сильно устойчива, то при $n\Delta t \leq T$

$$\|u^n\| \leq K_1^{1/2} \|u^n\|_H \leq (1 + K_2 \Delta t)^n K_1^{1/2} \|u^0\|_H \leq K_1 e^{K_2 T} \|u^0\|.$$

Это определение устойчивости по-прежнему весьма тесно связано с формой, в которой условие корректности применяется к задаче для дифференциального уравнения. Действительно, если построена норма, в которой решение этой задачи удовлетворяет условию (5.9), то при определенных условиях гладкости ее можно использовать и для разностной задачи. Однако эти условия гладкости трудно проверять, и одним из существенных достоинств изложенного в § 5.4 подхода Крайса к разностной задаче является то, что там эти трудности обходятся путем непосредственного исследования разностной задачи.

Наконец, остановимся вкратце на возможности использования в качестве основной не среднеквадратичной нормы, а какой-нибудь другой, в частности максимума модуля, нормы, очень привлекательной для вычислительных работ. В гл. 1 мы привели пример, показывающий, как она используется в одном простом случае, и можно было бы ожидать, что она особенно подходит для параболических уравнений, рассматриваемых в следующем параграфе. Однако получающееся при этом условие на множитель перехода может быть также применено к явной разностной схеме для любого уравнения вида $\partial u / \partial t = \partial^m u / \partial x^m$. Достаточность этого условия для устойчивости в норме максимум модуля была установлена Стрэнгом [1962], а необходимость позднее доказал Томи [1965a], который также обобщил эти результаты на неявные уравнения и на произвольную L_p -норму с $p \neq 2$ [1965b]. (L_p -норма $\|u\|_p$ определяется равенством $\|u\|_p^p = \int |u|^p dx$). Исходя из этого условия, Томи показывает, что при $m = 1$, т. е. в гиперболическом случае, любая устойчивая схема должна иметь нечетный порядок точности. Следовательно, все широко применяемые схемы второго порядка точности оказываются неустойчивыми во всех нормах, кроме среднеквадратичной. Возможный порядок роста весьма незначителен; Томи показывает, что если $\|C^n\| = 1$, то $\|C^n\|_p \leq (Kn)^{1/2 - 1/p}$, где K — некоторая постоянная. Но этого уже достаточно, чтобы считать эти нормы не подходящими для таких случаев.

§ 5.3. Параболические уравнения

Для параболического уравнения с одной пространственной переменной

$$\frac{\partial u}{\partial t} = a_0 \frac{\partial^2 u}{\partial x^2} + a_1 \frac{\partial u}{\partial x} + a_2 u, \quad a_0 > 0, \quad (5.10)$$

общая явная разностная схема имеет вид

$$u^{n+1}(x) = (Cu^n)(x) \equiv \sum_{\beta} c^{\beta} u^n(x + \beta \Delta x), \quad (5.11)$$

где суммирование ведется по некоторым фиксированным значениям β . Здесь a_i и c^{β} могут быть функциями от x , а c^{β} кроме того зависит от Δt и Δx .

Если представить $u^n(x + \beta\Delta x)$ в виде ряда Тэйлора, то условие согласованности приводит к соотношениям

$$\begin{aligned}\lim_{\Delta t \rightarrow 0} \frac{1}{\Delta t} \sum_{\beta} (c^{\beta} - \delta_0^{\beta}) &= a_2, \\ \lim_{\Delta t \rightarrow 0} \frac{\Delta x}{\Delta t} \sum_{\beta} \beta c^{\beta} &= a_1, \\ \lim_{\Delta t \rightarrow 0} \frac{(\Delta x)^2}{\Delta t} \sum_{\beta} \frac{1}{2} \beta^2 c^{\beta} &= a_0,\end{aligned}\tag{5.12}$$

где символ Кронекера $\delta_i^j = 0$ при $i \neq j$ и $= 1$ при $i = j$.

В случае когда коэффициенты a_i постоянны, устойчивость схемы (5.11) определяется множителем перехода

$$G = \sum_{\beta} c^{\beta} e^{i\beta\xi},\tag{5.13}$$

который в данном случае является тригонометрическим полиномом конечной степени от $\xi = k\Delta x$. Поэтому ясно, что для устойчивости коэффициенты c^{β} должны быть ограничены при $\Delta t \rightarrow 0$. Так как $a_0 > 0$, то в силу последнего из соотношений (5.12) $\Delta t = O((\Delta x)^2)$ при $\Delta t \rightarrow 0$. Положим $\Delta t/(\Delta x)^2 = \text{const}$; без потери общности эту постоянную можно считать равной единице. Наряду с этим предположим, что коэффициенты c^{β} могут быть представлены в виде

$$c^{\beta}(x, \Delta t) = c_0^{\beta}(x) + (\Delta t)^{1/2} c_1^{\beta}(x) + \Delta t c_2^{\beta}(x, \Delta t),\tag{5.14}$$

где все c_i^{β} равномерно ограничены и $c_2^{\beta}(x, \Delta t) \rightarrow c_2^{\beta}(x, 0)$ при $\Delta t \rightarrow 0$ равномерно по x . Тогда условие согласованности можно переписать в виде

$$\begin{aligned}\sum_{\beta} c_0^{\beta} &= 1, & \sum_{\beta} \beta c_0^{\beta} &= 0, & \sum_{\beta} c_1^{\beta} &= 0, \\ \sum_{\beta} \frac{1}{2} \beta^2 c_0^{\beta} &= a_0, & \sum_{\beta} \beta c_1^{\beta} &= a_1, & \sum_{\beta} c_2^{\beta} &= a_2.\end{aligned}\tag{5.15}$$

Прежде всего мы обоснуем сделанное ранее утверждение о том, что в большинстве случаев для задач с постоянными коэффициентами устойчивость определяется главной частью разностного оператора, т. е. значением G при $\Delta t = 0$. Подставляя разложения (5.14) в (5.13), получаем

$$G(\Delta t, \xi) = G_0(\xi) + (\Delta t)^{1/2} G_1(\xi) + \Delta t G_2(\Delta t, \xi),\tag{5.16}$$

где

$$G_j = \sum_{\beta} c_j^{\beta} e^{i\beta\xi}, \quad j = 0, 1, 2.$$

Следовательно, значение $G(0, \xi) = G_0(\xi)$ соответствует аппроксимации дифференциального оператора, у которого в силу (5.15) $a_1 = a_2 = 0$. Теперь предположим, что главная часть устойчива, т. е. что $|G_0(\xi)| \leq 1$. При доказательстве устойчивости схемы (5.11) член с G_2 очевидно несуществен. Используя (5.15), находим

$$G_0(\xi) = 1 - a_0 \xi^2 + O(\xi^3), \quad (5.17)$$

$$G_1(\xi) = ia_1 \xi + O(\xi^2). \quad (5.18)$$

Для наиболее простых аппроксимаций G_0 принимает действительные, а G_1 — мнимые значения; это справедливо также и во многих других случаях, и отсюда вытекает, что

$$|G_0 + (\Delta t)^{1/2} G_1| = (|G_0|^2 + \Delta t |G_1|^2)^{1/2} \leq 1 + O(\Delta t),$$

откуда и следует устойчивость. В общем случае, однако, для доказательства нашего утверждения мы должны предположить, что $|G_0(\xi)| < 1$ всюду на отрезке $[-\pi, \pi]$, за исключением точки $\xi = 0$. Тогда в силу (5.17)

$$|G_0(\xi)| \leq 1 - \delta \xi^2 \quad \text{для } |\xi| \leq \pi \quad (5.19)$$

при некотором $\delta > 0$. Обозначая действительную и мнимую части G_j через E_j и F_j соответственно, получаем

$$\begin{aligned} |G_0 + (\Delta t)^{1/2} G_1|^2 &= [E_0 + (\Delta t)^{1/2} E_1]^2 + [F_0 + (\Delta t)^{1/2} F_1]^2 = \\ &= |G_0|^2 + 2(\Delta t)^{1/2} (E_0 E_1 + F_0 F_1) + O(\Delta t). \end{aligned}$$

Но $E_0 E_1 + F_0 F_1 = O(\xi^2)$, и поэтому при достаточно малых Δt второй член не превосходит $\delta \xi^2$, откуда и следует устойчивость. Подробное обсуждение частных случаев этого результата проводится в § 8.4.

Стоит также отметить, что для любой разностной схемы, аппроксимирующей уравнение (5.10) с $a_1 = a_2 = 0$ и постоянным a_0 , как показал Стрэнг [1962], для устойчивости в норме максимум модуля множитель перехода должен удовлетворять одному из условий (1) $G(\xi) = ce^{i\beta\xi}$, где $|c| = 1$ и β действительно, или (2) $|G(\xi)| < 1$ при $|\xi| \leq \pi$, за исключением быть может конечного числа точек ξ_j , где $|G(\xi)| = 1$, причем для каждой такой точки существуют такие постоянные α_j , γ_j , ν_j , где α_j действительно, $\operatorname{Re} \gamma_j > 0$ и ν_j — четное натуральное число, что

$$G(\xi_j + \xi) = G(\xi_j) \exp[i\alpha_j \xi - \gamma_j \xi^{\nu_j} [1 + o(1)]] \quad \text{при } \xi \rightarrow 0.$$

Стрэнг доказал достаточность этих условий, а затем Томи [1965а] доказал их необходимость для устойчивости в норме максимум модуля. Поскольку условие (1) непременно к рассматриваемой нами задаче, то надо думать, что условие (5.19)

не является чрезмерно ограничительным даже для случая постоянных коэффициентов, если мы хотим иметь устойчивость не только в L_2 -норме, но и в норме максимум модуля. Более того, Видлунд [1966] дал пример неустойчивой разностной схемы для системы из двух параболических уравнений с переменными коэффициентами, где собственные значения главной части матрицы перехода удовлетворяют условию (2), которое не сводится к (5.19). Схема оказывается неустойчивой даже в L_2 -норме, так что (5.19), по-видимому, вообще наилучшее достаточное условие, на которое мы можем рассчитывать.

Возвращаясь к задаче с переменными коэффициентами, мы для простоты рассмотрим случай одного уравнения

$$\frac{\partial u}{\partial t} = a(x) \frac{\partial^2 u}{\partial x^2}, \quad a(x) \geq \alpha > 0,$$

которое аппроксимируется разностной схемой (5.11) с коэффициентами c^β , не зависящими от Δt . Все основные идеи отчетливо видны уже в этом случае. Хотя c^β теперь зависит от x , мы все же можем ввести локальный множитель перехода

$$G(x, \xi) = \sum_{\beta} c^\beta(x) e^{i\beta\xi} = 1 - a(x)\xi^2 + O(\xi^3), \quad (5.20)$$

который будем называть *символом оператора* S . Теперь мы докажем, что *если коэффициент $a(x)$ удовлетворяет условию Липшица и*

$$|G(x, \xi)| < 1 \text{ для всех } x \text{ и всех } \xi \neq 0 \text{ из интервала } [-\pi, \pi], \quad (5.21)$$

а $G(x, 0) = 1$ (согласованность), то разностная схема устойчива.

Из сделанных предположений вытекает, что при некотором $\delta > 0$

$$|G(x, \xi)|^2 \leq 1 - \delta\xi^2 \text{ для } |\xi| \leq \pi \text{ и всех } x. \quad (5.22)$$

Мы начнем с того, что рассмотрим устойчивость в окрестности произвольной точки, которую для удобства примем равной нулю. Итак, мы сосредоточим свое внимание на функциях u , для которых

$$u(x) \equiv 0 \text{ при } |x| \geq \xi > 0. \quad (5.23)$$

Обозначая через C_0 разностный оператор (с постоянными коэффициентами), получающийся при замене $c^\beta(x)$ в (5.11) на $c^\beta(0)$, а через $\hat{u}$ — преобразование Фурье от u , имеем

$$\begin{aligned} \|C_0 u\|^2 &= \int |G(0, \xi)|^2 |\hat{u}|^2 dk \leq \int (1 - \delta\xi^2) |\hat{u}|^2 dk \leq \\ &\leq \int [1 - \delta(2 \sin^{1/2} \xi)^2] |\hat{u}|^2 dk = \|u\|^2 - \delta \|u\|_1^2; \end{aligned} \quad (5.24)$$

здесь положено

$$\|u\|_p^2 = \int (2 \sin \frac{1}{2} \xi)^{2p} |\hat{u}|^2 dk = \|\Delta_+^p u\|^2 = \|\Delta_-^p u\|^2$$

(последние два равенства следуют из того, что преобразование Фурье от оператора разности вперед (назад) $\Delta_{\pm}$ равно $2ie^{\pm \frac{1}{2} i \xi} \sin \frac{1}{2} \xi$); ясно также, что

$$\|u\|_p \leq 2^{p-q} \|u\|_q \quad \text{при } p \geq q. \quad (5.25)$$

Поэтому для того, чтобы доказать устойчивость оператора C , положим

$$\|Cu\|^2 = \|C_0 u\|^2 + (\|Cu\|^2 - \|C_0 u\|^2)$$

и покажем, что последний член оценивается величиной $\delta \|u\|_1^2 + O(\Delta t) \|u\|^2$.

Используя согласованность в форме (5.20), напомним

$$C = I + a(x) \Delta_+ \Delta_- + Q, \quad (5.26)$$

где Q — разностный оператор по меньшей мере третьего порядка, равный сумме конечного числа членов, каждый из которых представляет собой произведение расположенных в определенном порядке разностных операторов Δ_+ или Δ_- и коэффициентов $a(x)$; например, следующим членом в этом представлении может быть $\frac{1}{2} a \Delta_+ \Delta_- a \Delta_+ \Delta_-$. Таким образом,

$$\begin{aligned} \|Cu\|^2 = & \|u\|^2 + 2 \operatorname{Re} (u, a \Delta_+ \Delta_- u) + \|a \Delta_+ \Delta_- u\|^2 + \\ & + 2 \operatorname{Re} (u, Qu) + 2 \operatorname{Re} (a \Delta_+ \Delta_- u, Qu) + \|Qu\|^2. \end{aligned} \quad (5.27)$$

Но из элементарного тождества $y_1 z_1 - y_2 z_2 \equiv y_1 (z_1 - z_2) + z_2 (y_1 - y_2)$ следует, что в выражении вида $(u, \Delta(a v))$ разностный оператор можно поставить после a за счет вычитания дополнительного члена, не превосходящего $L \Delta x \|u\| \cdot \|v\|$, где L — постоянная из условия Липшица для функции $a = a(x)$; этот член обозначим через $\varphi(u, v)$. Поэтому для второго члена из приведенного выше для $\|Cu\|^2$ представления мы имеем

$$\begin{aligned} (u, a \Delta_+ \Delta_- u) &= (u, \Delta_+ (a \Delta_- u)) + \varphi(u, \Delta_- u) = \\ &= -(\Delta_- u, a \Delta_- u) + \varphi(u, \Delta_- u) \end{aligned}$$

(справедливость последнего равенства устанавливается суммированием по частям, см. § 6.2). Если сравнить это с тем, что должно получиться для $\|C_0 u\|^2$, то для разности получится выражение

$$-(\Delta_- u, [a - a(0)] \Delta_- u) + \varphi(u, \Delta_- u).$$

Далее, поскольку $u(x)$ обращается в нуль при $|x| > \zeta$, то $\Delta_- u(x)$ равняется нулю при $|x| > 2\zeta$, если $\Delta x < \zeta$. Поэтому полученная выше разность не превосходит $2L\zeta \|u\|_1^2 + L \Delta x \|u\| \|u\|_1$,

Члены более общего вида можно оценить аналогично. А именно, рассмотрим разностный оператор общего вида $R = a_1 \Delta_1 a_2 \Delta_2 \dots a_r \Delta_r a_{r+1}$, где a_i — равномерно ограниченные и удовлетворяющие условию Липшица коэффициенты, а Δ_i — элементарные разностные операторы. Последовательно переставляя a_i с Δ_j , получим операторы

$$\begin{aligned} & \Delta_1 a_1 R_1, \quad \text{где } R_1 = a_2 \Delta_2 \dots \Delta_r a_{r+1}, \\ \text{или } & S_2 \Delta_2 a_3 R_2, \quad \text{где } S_2 = \Delta_1 a_1 a_2, \quad R_2 = \Delta_3 \dots a_{r+1}, \\ & S_3 \Delta_3 a_4 R_3, \quad \text{где } S_3 = \Delta_1 a_1 a_2 a_3 \Delta_2, \quad R_3 = \Delta_4 \dots a_{r+1}, \\ & \dots \dots \dots \\ & S_{n-1} \Delta_{n-1} a_n R_{n-1}, \quad \text{где } S_{n-1} = \Delta_1 a_1 \dots a_{n-2}, \quad R_{n-1} = \Delta_{n-1} a_n, \\ & R_n = \Delta_1 a_1 a_2 \dots a_{n-1} \Delta_{n-1} a_n. \end{aligned}$$

Следовательно,

$$\begin{aligned} |(u, Ru) - (u, R_n u)| & \leq |(u, Ru) - (u, \Delta_1 a_1 R_1 u)| + \\ & + |(u, S_2 \Delta_2 a_3 R_2 u) - (u, S_3 \Delta_3 a_4 R_3 u)| + \dots \\ & \dots + |(u, S_{n-1} \Delta_{n-1} a_n R_{n-1} u) - (u, R_n u)| \leq \\ & \leq L \Delta x \{ \|u\| \|R_1 u\| + \|S_2^* u\| \|R_2 u\| + \dots + \|S_{n-1}^* u\| \|R_{n-1} u\| \}, \end{aligned}$$

где L — постоянная Липшица для a_i , а S_i^* — оператор, сопряженный с S_i . Так как каждый из этих операторов ограничен, то выражение в фигурных скобках не превосходит $\text{const } \|u\|^2$, но здесь нам нужна более точная оценка. Каждый оператор S_i^* оканчивается оператором Δ_i , так что все члены, кроме первого, не превосходят $\text{const } \|u\| \|u\|$; если $r \geq 2$, то в результате перестановки двух последних множителей в операторе $R_1 \equiv S_1 \Delta_r a_{r+1}$ получим

$$|\|R_1 u\|^2 - \|S_1 a_{r+1} \Delta_r u\|^2| \leq L \Delta x \|S_1\| \|u\| (\|R_1 u\| + \|S_1 a_{r+1} \Delta_r u\|).$$

Следовательно,

$$\|R_1 u\| \leq \text{const} (\|u\|_1 + \Delta x \|u\|).$$

Поэтому мы можем сказать, что если оператор R имеет по крайней мере второй порядок, то

$$|(u, Ru) - (u, R_n u)| \leq \text{const } \Delta x \|u\| (\|u\|_1 + \Delta x \|u\|). \quad (5.28)$$

Приведя таким путем все члены в (5.27) к виду

$$(\Delta u, \left(\prod_i a_i \right) \left(\prod_j \Delta_j \right) u),$$

мы, очевидно, получим, что при $\Delta x < \zeta$

$$|\|Cu\|^2 - \|C_0 u\|^2| \leq M(\zeta) \|u\|_1^2 + K_1 \Delta x \|u\| \|u\|_1 + K_2 (\Delta x)^2 \|u\|^2,$$

где $M(\zeta) \rightarrow 0$ при $\zeta \rightarrow 0$, а K_1 и K_2 — некоторые постоянные. Возьмем ζ столь малое, скажем $\zeta = \zeta_0$, что $M(\zeta_0) < \delta/4$. Кроме того, мы имеем

$$K_1 \Delta x \|u\| \|u\|_1 \leq \frac{K_1^2 (\Delta x)^2}{\delta} \|u\|^2 + \frac{\delta}{4} \|u\|_1^2.$$

Сопоставляя эти оценки с (5.24), получаем, что при $\Delta t = (\Delta x)^2$ и $\Delta t < \zeta_0^2$

$$\|Cu\|^2 \leq [1 + (K_2 + K_1^2/\delta) \Delta t] \|u\|^2 - 1/2 \delta \|u\|_1^2, \quad (5.29)$$

откуда и следует устойчивость для функций u , удовлетворяющих условию (5.23).

Чтобы распространить этот результат на функции u общего вида, применим метод, предложенный Гордингом [1953]. Пусть $\sum_j d_j^2(x) \equiv 1$ есть гладкое разложение единицы, в котором каждый член $d_j(x)$ — гладкая функция от x , принимающая на некотором интервале I_j значения, заключенные между 0 и 1, и обращающаяся в нуль вне этого интервала, причем в каждой точке x лишь равномерно ограниченное число функций $d_j(x)$ отлично от нуля¹⁾. Каждый интервал I_j выберем столь малым, чтобы оценка (5.29) была справедлива для всех u , носители которых лежат в I_j ; при этом все длины I_j могут быть взяты меньшими, чем $2\zeta_0$ (величина ζ_0 была выбрана независимо от положения интервала и от Δt). Теперь мы можем, применив (5.29) к каждой функции $d_j(x)u(x)$ получить оценку суммы $\sum_j \|Cd_j u\|^2$ и сравнить это с $\|Cu\|^2 = \sum_j \|d_j Cu\|^2$. Для этого надо оценить результат перестановки разностных операторов с умножением на гладкие функции d_j . Рассуждая аналогично предыдущему, находим, что

$$\begin{aligned} |\|d_j u\|_1^2 - \|d_j \Delta_+ u\|^2| &= O(\Delta x) \|u_j\| (\|d_j \Delta_+ u\| + \|\Delta_+ d_j u\|) \leq \\ &\leq \frac{2}{3} (\|d_j \Delta_+ u\|^2 + \|\Delta_+ d_j u\|^2) + O(\Delta t) \|u\|^2, \end{aligned}$$

откуда

$$\|d_j u\|_1^2 \geq 1/2 \|d_j \Delta_+ u\|^2 + O(\Delta t) \|u_j\|^2.$$

В каждой точке x лишь конечное число функций d_j вносит вклад в последний член. Поэтому, меняя местами интегрирование и суммирование (после суммирования по j), получаем

$$\sum_j \|d_j u\|_1^2 \geq \frac{1}{2} \sum_j \|d_j \Delta_+ u\|^2 + O(\Delta t) \|u\|^2 = \frac{1}{2} \|u\|_1^2 + O(\Delta t) \|u\|^2.$$

¹⁾ По поводу построения такого разложения см., например, Гельфанд и Шилов [1959], стр. 182. — Прим. перев.

Следовательно, применяя к членам, возникающим при рассмотрении оператора C , аналогичные рассуждения и используя (5.29), мы получим

$$\begin{aligned} \|Cu\|^2 &= \sum_j \|d_j Cu\|^2 \leq \\ &\leq \sum_j \|Cd_j u\|^2 + O(\Delta x) \|u\| \|u\|_1 + O(\Delta t) \|u\|^2 \leq \\ &\leq [1 + O(\Delta t)] \sum_j \|d_j u\|^2 - \frac{1}{2}\delta \sum_j \|d_j u\|^2 + \\ &+ \frac{1}{4}\delta \|u\|_1^2 + O(\Delta t) \|u\|^2 \leq [1 + O(\Delta t)] \|u\|^2. \end{aligned}$$

Этим завершается доказательство теоремы.

Этот несколько ограниченный результат легко обобщается в различных направлениях. Например, в дифференциальное уравнение могут быть включены такие же младшие члены, как и в (5.10), для этого нужно лишь применить условие (5.21) к множителю перехода при $\Delta t = 0$, следуя рассуждениям, приведенным в начале этого параграфа. Такого рода результат был впервые доказан Фрицем Джоном [1952], использовавшим норму максимум модуля и предполагавшим, что функции a_0 , da_0/dx , $\partial^2 a_0/\partial x^2$, a_1 , da_1/dx и a_2 непрерывны и равномерно ограничены; помимо того, что Джон доказал устойчивость в этой норме, он не делал каких бы то ни было предположений о корректности постановки краевой задачи для дифференциального уравнения, а выводил существование, единственность и другие свойства решения из свойств разностного уравнения. Он пошел даже еще дальше, заменив в (5.10) член $a_2 u$ нелинейным членом $d(x, t, u)$, равномерно удовлетворяющим условию Липшица по u .

В следующем параграфе будет показано, что обобщение на случай нескольких пространственных переменных осуществляется непосредственно. Этот результат содержится в статье Лакса и Вендроффа [1962], которым принадлежат основные идеи изложенного нами подхода к задачам с переменными коэффициентами. Однако в деталях наше изложение ближе к трактовке Крайса [1964] и Видлунда [1965]; последнему принадлежит дальнейшее обобщение результатов на системы параболических уравнений общего вида, для решения которых используется широкий класс неявных многослойных методов. Впоследствии Видлунд [1966] распространил эти общие теоремы на случай нормы максимум модуля, обобщив таким образом результаты Джона, Аронсона [1963] и других. Во всех случаях существенную роль играет условие (5.22), применен-

ное к собственным значениям символа главной части разностного оператора, поэтому Видлунд называет разностные схемы, удовлетворяющие этому условию, параболическими.

§ 5.4. Диссипативные разностные схемы для симметричных гиперболических уравнений

Рассмотрим краевую задачу для гиперболической системы

$$\frac{\partial \mathbf{u}}{\partial t} = \sum_{j=1}^d A_j(\mathbf{x}) \frac{\partial \mathbf{u}}{\partial x_j}, \quad -\infty < x_j < +\infty, \quad 0 \leq t \leq 1, \quad (5.30)$$

где вектор $\mathbf{u} = \mathbf{u}(\mathbf{x}, t)$ имеет p компонент и $p \times p$ -матрицы $A_j(\mathbf{x})$ гладко зависят от $\mathbf{x}$, но не зависят от t . Будем предполагать, если специально не оговорено противное, что матрицы A_j эрмитовы. Введем матрицу $P(\mathbf{x}, \mathbf{k})$, где

$$P(\mathbf{x}, \mathbf{k}) = \sum_{j=1}^d A_j(\mathbf{x}) k_j, \quad -\infty < k_j < +\infty; \quad (5.31)$$

в случае постоянных коэффициентов эта матрица представляет собой преобразование Фурье от оператора из правой части уравнения (5.30) (см. § 4.3). Предположение о гиперболичности состоит в том, что существует такая невырожденная матрица $T = T(\mathbf{x}, \mathbf{k})$, что

$$T P T^{-1} = \begin{bmatrix} \mu_1 & & & \\ & \mu_2 & & \\ & & \ddots & 0 \\ 0 & & & \ddots \\ & & & & \mu_p \end{bmatrix}, \quad (5.32)$$

где $\mu_\nu = \mu_\nu(\mathbf{x}, \mathbf{k})$ действительны, а T и T^{-1} равномерно ограничены по $\mathbf{x}$ и $\mathbf{k}$, т. е.

$$\|T(\mathbf{x}, \mathbf{k})\|, \quad \|T^{-1}(\mathbf{x}, \mathbf{k})\| \leq K_0.$$

Если матрицы A_j эрмитовы, то матрица T унитарна, и указанные условия очевидным образом выполнены.

Будем считать для удобства, что в аппроксимирующих разностных схемах $\Delta x_j = \Delta t$ при всех j . Кроме того, будем рассматривать только явные схемы

$$\mathbf{u}^{n+1} = C(\Delta t) \mathbf{u}^n, \quad (5.33)$$

где $C(\Delta t)$ — конечная сумма операторов сдвига с матричными коэффициентами:

$$(C(\Delta t)u)(x) \equiv \sum_{\beta} c^{\beta}(x, \Delta t) u(x + \beta \Delta t). \quad (5.34)$$

Будем предполагать, что матрицы $c^{\beta}(x, \Delta t)$ эрмитовы, если эрмитовы $A_j(x)$. Локальная матрица перехода, или символ G , имеет вид

$$G(x, \Delta t, \xi) = \sum_{\beta} c^{\beta}(x, \Delta t) e^{i\beta \cdot \xi}, \quad (5.35)$$

где $\xi = \Delta tk$. Обозначим собственные значения матрицы G через λ_v и введем следующее определение.

Разностная схема (5.33) называется *диссипативной* порядка $2r$, где r — натуральное число, если существует такое $\delta > 0$, что для всех ξ , удовлетворяющих условию $\max |\xi_j| \leq \pi$, всех x и всех Δt , меньших некоторого $\tau > 0$,

$$|\lambda_v(x, \Delta t, \xi)| \leq 1 - \delta |\xi|^{2r}, \quad (5.36)$$

где $|\xi|$ — евклидова норма вектора ξ .

Вспоминая определение порядка точности (4.12а) или (4.13), мы можем сформулировать следующую теорему, принадлежащую Крайсу [1964].

Предположим, что матрицы, входящие в дифференциальное уравнение (5.30) и в аппроксимирующее его разностное уравнение (5.33), эрмитовы, равномерно ограничены и равномерно удовлетворяют условию Липшица по x . Тогда, если схема (5.33) является диссипативной порядка $2r$ и имеет порядок точности $2r - 1$ при некотором натуральном r , то она устойчива.

Доказательство теоремы весьма длинно, но в общем естественно; некоторые упрощения в первоначальное доказательство Крайса были внесены Парлеттом [1966]. Имеются три основных этапа. Сначала мы покажем, что сделанные предположения достаточны для устойчивости в случае постоянных коэффициентов, т. е. достаточны для локальной устойчивости. Как мы видели в § 4.9, это равносильно тому, что для каждого фиксированного x существует такая положительно определенная эрмитова матрица $\hat{H} = \hat{H}(x, \Delta t, \xi)$, что $G^* \hat{H} G \leq \hat{H}$ и для некоторой константы $K_2 > 0$

$$K_2^{-1} I \leq \hat{H} \leq K_2 I. \quad (5.37)$$

Если бы $\hat{H}$ достаточно гладко зависело от x , то последующие рассуждения соответственно упростились бы (см. следующий параграф), но чаще всего это не так и во всяком случае это

трудно доказать. Мы построим матрицу $\hat{H}$ так, что она будет пригодна для целой окрестности точки x ; она будет иметь период 2π по каждой компоненте вектора ξ , и нам легко будет удовлетворить неравенству

$$G^* \hat{H} G \leq (1 - 1/2\delta) |\xi|^{2r} \hat{H}, \quad (5.38)$$

но нам потребуется дополнительно, чтобы матрица $\hat{H}$ имела специальный вид

$$\hat{H} = I + \hat{H}_{2r-1}(x, \Delta t, \xi), \text{ где } \|\hat{H}_{2r-1}\| \leq \text{const} |\xi|^{2r-1}, \max_j |\xi_j| \leq \pi. \quad (5.39)$$

Заключительный этап доказательства состоит в том, что с помощью $\hat{H}$ строится норма $\|\cdot\|_H$, в которой решение уравнения (5.33) удовлетворяет сильному условию устойчивости

$$\|u^{n+1}\|_H^2 \leq [1 + O(\Delta t)] \|u^n\|_H^2. \quad (5.40)$$

Заметим прежде всего, что зависимость C и тем самым G и $\hat{H}$ от Δt осуществляется через матрицы $c^p(x, \Delta t)$, вычисленные для значений A_j в точках, соседних с x . Так как A_j удовлетворяют условию Липшица, то величина $\|C(\Delta t) - C(0)\|(\Delta t)^{-1}$ ограничена и потому свойства устойчивости оператора $C(\Delta t)$ и его главной части $C(0)$ одинаковы. Поэтому мы рассмотрим только эту последнюю и будем считать в дальнейшем, что C , G , $\hat{H}$ и λ не зависят от Δt .

На первом этапе доказательства мы опустим предположение о том, что матрицы A_j эрмитовы. Это уведет нас несколько в сторону от намеченной цели, но будет полезно для выяснения роли различных предположений. В силу двойственного определения порядка точности (4.13) имеем

$$G(\xi) = e^{iP(\xi)} + Q(\xi), \quad (5.41)$$

где $\|Q\| \leq \text{const} |\xi|^{2r}$ при $\max_j |\xi_j| \leq \pi$. Далее, в силу гиперболичности это соотношение можно с помощью матрицы T из (5.32) преобразовать к виду $TGT^{-1} = M + R$, где матрица $M = \exp(iTP T^{-1})$ диагональна и $\|P\| \leq \text{const} |\xi|^{2r}$. Если теперь перейти к верхней треугольной форме с помощью унитарного преобразования U , то мы получим

$$\tilde{G} = UTGT^{-1}U^* = UMU^* + URU^* = \Lambda + N, \quad (5.42)$$

где Λ — диагональная матрица из собственных значений λ_j матрицы G , а N — нильпотентная матрица, у которой $N_{ij} = 0$ при $i \geq j$, причем важно отметить, что $\|N\| \leq K_1 |\xi|^{2r}$ для некоторой постоянной $K_1 > 1$. Это устанавливается следующим образом. Элементы матрицы URU^* ограничены скалярным

кратным величины $|\xi|^{2r}$, и поскольку $\tilde{G}$ имеет верхнюю треугольную форму, то так же ограничены элементы матрицы UMU^* , лежащие ниже диагонали. Но UMU^* — нормальная матрица; и потому ее наддиагональные элементы ограничены скалярным кратным оценки для поддиагональных элементов¹⁾, откуда и следует наше утверждение.

Теперь непосредственной проверкой легко убедиться в том, что степени G ограничены. Действительно, как и в § 5.2,

$$\begin{aligned} \|G^n\| &\leq K_0^2 \|\tilde{G}^n\| \leq K_0^2 \sum_{j=0}^{p-1} \binom{n}{j} \|\Lambda\|^{n-j} \|N\|^j \leq \\ &\leq K_0^2 (K_1/\delta)^{p-1} \sum_{j=0}^{p-1} \binom{n}{j} (1 - \delta|\xi|^{2r})^{n-j} (\delta|\xi|^{2r})^j \leq \\ &\leq K_0^2 (K_1/\delta)^{p-1} \quad \text{при } n \geq p-1. \end{aligned}$$

Менее очевидным является построение матрицы $\hat{H} = \hat{H}(\xi)$, аналогичное тому построению, которое проводилось при доказательстве матричной теоремы Крайса в предыдущей главе. Вводя диагональную матрицу

$$D = \begin{bmatrix} \Delta & & & 0 \\ & \Delta^2 & & \\ & & \ddots & \\ 0 & & & \Delta^p \end{bmatrix}$$

с $\Delta > 1$, выберем Δ столь большим, чтобы

$$\left(1 - \frac{1}{2} \delta |\xi|^{2r}\right) D - \tilde{G}^* D \tilde{G} \geq 0,$$

т. е.

$$\left(1 - \frac{1}{2} \delta |\xi|^{2r}\right) I - \tilde{G}_D^* \tilde{G}_D \geq 0,$$

¹⁾ Имеем

$$A^* A = A A^* \rightarrow \sum_{k=1}^p |a_{ik}|^2 = \sum_{k=1}^p |a_{ki}|^2.$$

Предположим, что $|a_{ik}| \leq K$ при $i > k$. Тогда

$$\sum_{k=2}^p |a_{1k}|^2 \leq (p-1) K^2, \quad \sum_{k=3}^p |a_{2k}|^2 \leq (p-2) K^2 + |a_{12}|^2 \leq (2p-3) K^2$$

и т. д. По индукции находим, что при $i < k$

$$|a_{ik}|^2 \leq [1 + (p-2) 2^{i-1}] K^2.$$

где через R_D обозначено $D^{1/2}RD^{-1/2}$. Ясно, что это возможно, так как в силу (5.42) $\tilde{G}_D = (\Lambda + N)_D = \Lambda + N_D$, а поскольку (j, k) -элемент матрицы N_D равен $\Delta^{1/2(j-k)}N_{jk}$, то $\|N_D\| \leq \Delta^{-1/2}\|N\| \leq K_1\Delta^{-1/2}|\xi|^{2r}$. Следовательно,

$$\begin{aligned} \|\tilde{G}_D^*\tilde{G}_D\| &\leq \|\Lambda^*\Lambda\| + 2\|N_D\| + \|N_D\|^2 \leq \\ &\leq 1 - \delta|\xi|^{2r} + \text{const}\Delta^{-1/2}|\xi|^{2r} \leq 1 - 1/2\delta|\xi|^{2r} \end{aligned}$$

для достаточно больших Δ . Таким образом, для достаточно больших Δ неравенство (5.38) удовлетворяется при $\hat{H} = T^*U^*DUT$, причем $\hat{H}$ удовлетворяет соотношению (5.37) при $K_2 \leq \Delta^p K_0^2$.

Чтобы построить H с дополнительным свойством (5.39), применим другой подход. Будем считать матрицы A_j и c^b эрмитовыми и предположим, что матрица P предварительно приведена к диагональному виду при помощи унитарного преобразования. Положим также $\xi = \sigma\xi_0$, где ξ_0 — некоторый фиксированный единичный вектор, так что

$$G = M + \sigma^{2r}R + o(\sigma^{2r}), \quad (5.43)$$

где матрица $M = \exp(i\sigma P(\xi_0))$ диагональна и унитарна, а матрица $R = R(\xi_0)$ эрмитова. Теперь предположим, что два собственных значения μ_j и μ_k матрицы P равны. Тогда имеется некоторый произвол в выборе унитарного преобразования, диагонализующего P , а именно это преобразование определяется с точностью до любого унитарного преобразования в двумерном подпространстве, отвечающем этим двум собственным значениям. Так как матрица R эрмитова, то указанный произвол можно использовать для обращения в нуль элемента R_{jk} .

Предполагая это выполненным, определим с помощью R эрмитову матрицу $\hat{H}'$ формулой

$$\hat{H}'_{jk} = \begin{cases} 0 & \text{при } \mu_j = \mu_k, \\ -2iR_{jk}/(\mu_j - \mu_k) & \text{при } \mu_j \neq \mu_k \end{cases} \quad (5.44)$$

и затем построим эрмитову матрицу

$$\hat{H} = I + \sigma^{2r-1}\hat{H}'. \quad (5.45)$$

Проверим выполнение условия (5.38). Разлагая M по степеням σ , получаем

$$G^*\hat{H}G = \hat{H} + \sigma^{2r}[2R + i(\hat{H}'P - P\hat{H}')] + o(\sigma^{2r}). \quad (5.46)$$

В силу определения $\hat{H}'$ внедиагональные элементы матрицы в квадратных скобках равны нулю. Кроме того, справедлива

следующая лемма, доказательство которой будет приведено несколько позже.

Лемма. Если матрица G диссипативной разностной схемы имеет вид (5.43), то

$$R_{jj} \leq -\delta + o(1) \quad \text{при } \sigma \rightarrow 0. \quad (5.47)$$

Поэтому для $\xi = \sigma\xi_0$ имеем

$$2R + i(\hat{H}'P - P\hat{H}') \leq -2\delta I + o(1),$$

причем матрицы $\hat{H}' = \hat{H}'(\xi_0)$ ограничены, хотя и не равномерно по ξ_0 . Таким образом, для каждого направления ξ_0 существует целый отрезок $0 \leq \sigma \leq \sigma_0(\xi_0)$, на котором выполнены соотношения (5.38) и (5.39). Далее, так как P и R непрерывно зависят от ξ , то найдется такое $\varepsilon > 0$, что при $|\xi_0 - \xi_1| < \varepsilon$ и $|\xi_1| = 1$

$$2R(\xi_1) + i[\hat{H}'(\xi_0)P(\xi_1) - P(\xi_1)\hat{H}'(\xi_0)] \leq -\delta I;$$

поэтому при $\xi = \sigma\xi_1$, $0 \leq \sigma \leq \sigma_1(\xi_0)$ условия (5.38) и (5.39) также выполнены. Но по лемме Гейне — Бореля весь шар $|\xi| \leq 1$ можно покрыть конечным числом таких окрестностей, в каждой из которых $\hat{H}'$ можно считать постоянной. Поэтому мы можем построить $\hat{H}(\xi)$ в некоторой окрестности нуля $|\xi| \leq \leq \sigma_2$ так, чтобы выполнялись все три требования (5.37) — (5.39). Вне этой окрестности соотношение (5.39) не имеет места, и можно использовать только матрицу $\hat{H}$, построенную на первом этапе.

Для доказательства леммы покажем, что при $\sigma \rightarrow 0$ собственные значения матрицы G равны ее диагональным элементам с точностью до членов порядка σ^{2r} . Это следует из уточненной теоремы Гершгорина (см. примечание на стр. 85), утверждающей, что если s гершгоринских кругов матрицы M отделены от остальных, то их объединение содержит в точности s собственных значений M ; эта теорема легко следует из соображений непрерывности, так как собственные значения матрицы непрерывно зависят от ее элементов. Предположим теперь, что μ_k — простое собственное значение матрицы P , и умножим k -ю строку матрицы G на $\sigma^{1/2}$, а ее k -столбец разделим на $\sigma^{1/2}$. Это преобразование подобия не изменяет собственных значений λ_j матрицы G , но делает радиус k -го гершгоринского круга равным $O(\sigma^{2r+1/2})$ а остальные радиусы равными $O(\sigma^{2r-1/2})$. Так как центрами этих кругов являются точки $G_{jj} = e^{i\sigma\mu_j} + \sigma^{2r}R_{jj} + O(\sigma^{2r+1})$ и $\mu_k \neq \mu_j$ при $j \neq k$, то k -й круг отделен от остальных при достаточно малых σ . Поэтому матрица G имеет собственное

значение

$$\lambda = e^{i\sigma\mu_k} + \sigma^{2r} R_{kk} + o(\sigma^{2r}).$$

Но в силу (5.36) $|\lambda| \leq 1 - \delta\sigma^{2r}$, так что

$$1 - 2\delta\sigma^{2r} \geq 1 + 2\sigma^{2r} R_{kk} + o(\sigma^{2r}),$$

т. е. $R_{kk} \leq -\delta + o(1)$ при $\sigma \rightarrow 0$.

Если μ_k не является простым собственным значением P , то все строки и столбцы G , соответствующие μ_k , преобразуем так же, как и выше. Так как $R_{jk} = 0$ при $\mu_j = \mu_k$ и $j \neq k$, то радиусы соответствующих кругов снова равны $O(\sigma^{2r+1/2})$. Поэтому по крайней мере один из них содержит собственное значение матрицы G , так что (5.47) выполняется для соответствующего R_{jj} . С другой стороны, если бы для некоторых из оставшихся кругов $R_{jj} > -\delta + o(1)$, то они были бы отделены от тех кругов, для которых $R_{jj} \leq -\delta + o(1)$ при $\sigma \rightarrow 0$, и мы пришли бы к противоречию с тем фактом, что тогда они содержали бы собственное значение G .

Переходя к заключительному этапу доказательства нашей основной теоремы, вернемся к рассмотрению задачи с переменными коэффициентами, определяемой уравнениями (5.30), (5.33) и (5.34), с эрмитовыми матрицами A_j и c^B . Эта часть доказательства по форме аналогична тому доказательству, которое было проведено для более простой задачи, рассмотренной в предыдущем параграфе, за исключением того, что матрица $\hat{H}$ не равна теперь тождественно единичной матрице. Начнем снова с рассмотрения устойчивости в окрестности произвольной точки, которую для удобства будем считать нулем, и сперва ограничимся такими функциями u , для которых

$$u(x) \equiv 0 \quad \text{при } |x| \geq \xi > 0. \quad (5.48)$$

В этом случае разностному оператору S_0 с коэффициентами $c^B(0)$ соответствует локальная матрица перехода $G = G(0, \xi)$, получаемая с помощью (5.35). Далее, как мы уже показали, существует эрмитова матрица $\hat{H}(0, \xi)$ вида

$$\hat{H} = I + \hat{H}_{2r-1}(0, \xi) \quad \text{с} \quad \|\hat{H}_{2r-1}\| \leq K_3 |\xi|^{2r-1}, \quad (5.49)$$

удовлетворяющая условиям (5.37) и (5.38). С ее помощью определим оператор H следующим образом:

$$Hu(x) = (2\pi)^{-d/2} \int \hat{H}(0, \Delta tk) \hat{u}(k) e^{ik \cdot x} dk, \quad (5.50)$$

где $\hat{u}$ — преобразование Фурье функции u . Аналогично определим H_{2r-1} и обозначим $\{(u, Hu)\}^{1/2}$ через $\|u\|_H$. Очевидно, что

в силу свойств $\hat{H}$ эта величина действительно будет нормой, и

$$\begin{aligned}\|Cu\|_H^2 &= (Cu, [I + H_{2r-1}]Cu) = E + F = \\ &= (E_0 + F_0) + (E - E_0) + (F - F_0),\end{aligned}\quad (5.51)$$

где

$$E = \|Cu\|^2, \quad F = (Cu, H_{2r-1}Cu),$$

а E_0 и F_0 получаются из них заменой C на C_0 . Наша основная цель — оценить каждый из трех членов в правой части (5.51), чтобы установить неравенство (5.40). Имеем

$$\begin{aligned}E_0 + F_0 &= \|C_0u\|_H^2 = \int \hat{u}^* G^* \hat{H} G u \, dk \leq \\ &\leq \int \hat{u}^* (1 - 1/2 \delta |\xi|^{2r}) \hat{H} \hat{u} \, dk \leq \\ &\leq \int \hat{u}^* \left(1 - 1/2 \delta \sum_{j=1}^d (2 \sin 1/2 \xi_j)^{2r} \right) \hat{H} \hat{u} \, dk = \\ &= \|u\|_H^2 - 1/2 \delta K_2^{-1} \|u\|_r^2,\end{aligned}\quad (5.52)$$

где, как и в § 5.3,

$$\begin{aligned}\|u\|_p^2 &= \int \sum_{j=1}^d (2 \sin 1/2 \xi_j)^{2p} |\hat{u}|^2 \, dk = \\ &= \sum_{j=1}^d \|\Delta_{+j}^p u\|^2 = \sum_{j=1}^d \|\Delta_{-j}^p u\|^2,\end{aligned}\quad (5.53)$$

так что

$$\|u\|_p \leq 2^{p-q} \|u\|_q \quad \text{при } p \geq q. \quad (5.54)$$

Для того чтобы оценить остальные члены, необходимо подробнее рассмотреть структуру оператора C . Так как этот оператор является представлением дифференциального оператора с точностью до членов порядка $2r - 1$, то он однозначно определяется с точностью до членов этого порядка. Следовательно, мы можем записать (ср. с выражением (5.41) для матрицы G)

$$C = \sum_{m=0}^{2r-1} \frac{1}{m!} P_\Delta^m + Q_\Delta, \quad (5.55)$$

где

$$P_\Delta = \sum_{j=1}^d A_j(\mathbf{x}) \Delta_j^{(2r-1)},$$

а $\Delta_j^{(2r-1)}$ обозначает усеченное разложение оператора $\Delta t (\partial/\partial x_j)$ по степеням оператора центральной разности $\Delta_{0j} = \frac{1}{2} (\Delta_{+j} + \Delta_{-j})$:

$$\Delta_j^{(2r-1)} = \sum_{m=0}^{r-1} (-1)^m \gamma_{2m+1} \Delta_{0j}^{2m+1}, \quad (5.56)$$

где γ_{2m+1} — некоторые действительные постоянные. Остаточный член Q_Δ также есть конечная сумма произведений $\prod_i A \Delta^{p_i}$ разностных операторов, где каждое A представляет собой одну из матриц $A_j(x)$, а каждое Δ — один из операторов Δ_{+j} или Δ_{-j} ; при этом $2r \leq \sum_i p_i \leq \text{const}$, так как наша разностная схема по предположению является явной схемой $(2r-1)$ -го порядка точности.

Как уже отмечалось, в выражении $(u, \Delta A v)$ разностный оператор можно переставить с A за счет добавления дополнительного слагаемого, не превосходящего по норме $L \Delta t \|u\| \cdot \|v\|$, где L — постоянная Липшица для A ; в дальнейшем мы будем обозначать это слагаемое через $\varphi(u, v)$. Поэтому, используя простейшую формулу суммирования по частям $(u, \Delta_0 A v) = -(\Delta_0 u, A v)$ и тот факт, что матрица A эрмитова, получаем

$$(u, A \Delta_0 v) = -(A \Delta_0 u, v) + \varphi(u, v). \quad (5.57)$$

Аналогичные формулы справедливы для каждого из конечного числа разностных операторов нечетного порядка, образующих P_Δ , и тем самым для самого P_Δ . Поэтому

$$\begin{aligned} \left\| \left(\sum_{m=0}^{2r-1} \frac{1}{m!} P_\Delta^m \right) u \right\|^2 &= \left(u, \sum_{l=0}^{2r-1} \frac{1}{l!} (-P_\Delta)^l \left(\sum_{m=0}^{2r-1} \frac{1}{m!} P_\Delta^m \right) u \right) + \varphi(u, u) = \\ &= \|u\|^2 + (u, R u) + \varphi(u, u). \end{aligned} \quad (5.58)$$

Из этого выражения видно, что R является разностным оператором того же вида, что и Q_Δ , т. е. порядка не меньшего, чем $2r$. Чтобы оценить члены, содержащие такие операторы, будем рассуждать так же, как и при получении неравенства (5.28) в § 5.3: переставляя некоторые A и Δ , получаем

$$(u, R v) = (\Delta^r u, S \Delta^r v) + \varphi(u, v),$$

где S — ограниченный разностный оператор порядка на $2r$ меньшего, чем R , а Δ^r обозначает произведение r множителей $\Delta_{\pm j}$.

Следовательно,

$$\begin{aligned} \|Cu\|^2 &= \left\| \left(\sum_{m=0}^{2r-1} \frac{1}{m!} P_{\Delta}^m \right) u \right\|^2 + 2 \operatorname{Re} \left(Q_{\Delta} u, \left(\sum_{m=0}^{2r-1} \frac{1}{m!} P_{\Delta}^m \right) u \right) + \|Q_{\Delta} u\|^2 = \\ &= \|u\|^2 + (\Delta^2 u, S' \Delta^2 u) + \varphi(u, u), \end{aligned} \quad (5.59)$$

где S' , как и S , является ограниченным разностным оператором.

Но $E - E_0 = \|Cu\|^2 - \|C_0 u\|^2$, так что главным в этой разности будет член, соответствующий оператору $S' - S'_0$. Поскольку носитель функции u содержится в шаре радиуса ζ , то C можно заменить на $C_0 + g(x)(C - C_0)$, где $g(x)$ — гладкая функция, удовлетворяющая условиям $g(x) \equiv 1$ для $|x| \leq \frac{3}{2}\zeta$ и $g(x) \equiv 0$ для $|x| > 2\zeta$ при условии, что Δx достаточно мало. Поэтому для вычисления разности $\|Cu\|^2 - \|C_0 u\|^2$ при достаточно малых Δx достаточно интегрировать по шару радиуса 2ζ ; следовательно,

$$|E - E_0| \leq M_1(\zeta) \|u\|_r^2 + K_4 \Delta t \|u\|^2, \quad (5.60)$$

где $M_1(\zeta) \rightarrow 0$ при $\zeta \rightarrow 0$, а K_4 — некоторая постоянная.

Для оценки последнего члена используем специальный вид матрицы H (см. (5.49)):

$$\begin{aligned} |F - F_0| &= |\operatorname{Re}([C + C_0]u, H_{2r-1}[C - C_0]u)| \leq \\ &\leq |([C + C_0]u, H_{2r-1}[C - C_0]u)| \end{aligned} \quad (5.61)$$

(оператор C ограничен). Теперь представим H_{2r-1} в виде $H_{2r-1} = H_1 H_2$, где H_1 и H_2 определяются преобразованиями Фурье

$$\hat{H}_1 = \left(\sum_{j=1}^d (2 \sin^{1/2} \xi_j)^2 \right)^{r/2} I, \quad \hat{H}_2 = \hat{H}_1^{-1} \hat{H}_{2r-1}.$$

Имеем

$$|\xi|^r I \geq \hat{H}_1 \geq \left(\frac{2}{\pi} \right)^r |\xi|^r I \quad \text{и} \quad \|\hat{H}_2\| \leq \left(\frac{\pi}{2} \right)^r K_3 |\xi|^{r-1}.$$

Следовательно,

$$\begin{aligned} |[C + C_0]u, H_{2r-1}[C - C_0]u| &\leq \|H_1[C + C_0]u\| \|H_2[C - C_0]u\| \leq \\ &\leq \|[C + C_0]u\|_r \|[C - C_0]u\|_{r-1} \leq \operatorname{const} \|u\|_r \|[C - C_0]u\|_{r-1}. \end{aligned}$$

Рассуждая обычным образом и используя (5.53) и (5.54), получаем

$$\begin{aligned} \|[C - C_0]u\|_{r-1} &\leq \sum_{j=1}^d \|\Delta_{+j}^{r-1} [C - C_0]u\| \leq \\ &\leq \sum_{j=1}^d \|[C - C_0] \Delta_{+j}^{r-1} u\| + \operatorname{const} \Delta t \cdot \|u\| \leq \\ &\leq M_2(\zeta) \|u\|_r + K_5 \Delta t \|u\|, \end{aligned}$$

где $M_2(\zeta) \rightarrow 0$ при $\zeta \rightarrow 0$, так как оператор $C - C_0$ имеет по меньшей мере первый порядок по ζ . Применяя (5.60) и еще раз (5.54), находим, что

$$|E - E_0| + |F - F_0| \leq M(\zeta) \|u\|_r^2 + K_6 \Delta t \|u\|^2,$$

где $M(\zeta) \rightarrow 0$ при $\zeta \rightarrow 0$, а K_6 — некоторая постоянная. Таким образом, при $\zeta \leq \zeta_0$ для некоторого $\zeta_0 > 0$ первый член в правой части будет меньше вычитаемого из правой части (5.52); поэтому

$$\|u^{n+1}\|_H^2 \leq \|u^n\|_H^2 + \alpha K_2^{-1} \Delta t \|u^n\|^2 \leq (1 + \alpha \Delta t) \|u^n\|_H^2, \quad (5.62)$$

где $\alpha = K_2 K_6$.

Чтобы получить этот результат для функций u^n общего вида, мы снова применим метод Гординга [1953]. Пусть $\sum_i d_i^2(x) \equiv 1$ — гладкое разложение единицы, в котором каждый член $d_i(x)$ есть гладкая неотрицательная функция от x , равная нулю вне некоторого шара S_i , причем в каждой фиксированной точке x лишь равномерно ограниченное число членов отлично от нуля. Возьмем каждый шар S_i столь малым, чтобы оценка (5.62) выполнялась для всех функций u^n , носители которых принадлежат S_i ; из предыдущего ясно, что радиусы шаров S_i можно считать большими, чем ζ_0 . Теперь определим оператор H_i в центре шара S_i , как и выше, и положим

$$\|u\|_H^2 = \sum_i \|d_i u\|_{H_i}^2 = \sum_i (d_i u, H_i d_i u);$$

ясно, что эта норма по-прежнему удовлетворяет неравенствам, общим для всех H_i . Далее,

$$\|u^{n+1}\|_H^2 = \sum_i (d_i C u^n, H_i d_i C u^n).$$

Замечая, что в каждой точке x при перестановке d_i и C лишь конечное число функций d_i даст ненулевой вклад, находим

$$\begin{aligned} \|u^{n+1}\|_H^2 &\leq \sum_i (C d_i u^n, H_i C d_i u^n) + \text{const } \Delta t \|u^n\|^2 \leq \\ &\leq \sum_i \|d_i u^n\|_{H_i}^2 (1 + \alpha \Delta t) + \text{const } \Delta t \|u^n\|^2 \leq \\ &\leq \|u^n\|_H^2 [1 + O(\Delta t)]; \end{aligned}$$

тем самым устойчивость разностной схемы установлена, чем и завершается доказательство теоремы Крайса.

Мы привели столь подробное доказательство как по причине важности результата, так и ввиду общей ценности проведенных

рассуждений. Всегда полезно иметь результат, зависящий только от собственных значений матрицы перехода или символа G , но особенно важным оказалось понятие диссипативной разностной схемы для нелинейных задач (см., например, Рихтмаейр и Мортон [1964]). Важно еще и то, что тем же путем можно получить целый ряд аналогичных результатов. В частности, нашу теорему нельзя непосредственно применить к схемам Лакса — Вендроффа (см. § 12.7), имеющим четвертый порядок диссипативности, но только второй порядок точности. Этот и ряд других случаев можно свести к предыдущему заменой переменных ξ . А именно предположим, что

$$G(\xi) = \exp \{iP[g(\xi)]\} + Q(\xi),$$

где $\|Q\| \leq \text{const} |\xi|^{2r}$, а векторная функция $g(\xi)$ аналитична и такова, что $g_j = \xi_j + o(|\xi|)$ при $\xi \rightarrow 0$. (Для схем Лакса — Вендроффа $g_j = \sin \xi_j$.) Тогда переход к переменным g дает возможность применить изложенные выше соображения. Нетрудно убедиться, что для символа G такого вида разностная схема может иметь только первый порядок точности. Оказывается, что в общем случае, однако, условие порядка точности $2r - 1$ не может быть ослаблено такими методами, хотя Парлетт [1966] и доказал, что порядка $2r - 2$ достаточно, если P имеет различные собственные значения, т. е. если система дифференциальных уравнений строго гиперболична.

§ 5.5. Дальнейшие результаты для симметричных гиперболических уравнений

Имеется ряд теорем о разностных схемах для гиперболических уравнений, аналогичных теореме, доказанной в предыдущем параграфе. Если сделать о символе G предположения, в некотором смысле более сильные (хотя и не включающие в себя никаких условий диссипативности), то нет нужды в столь тщательном анализе информации о разностной схеме. Тогда оказывается более удобным использовать операторы сдвига $T^\beta u(x) = u(x + \beta \Delta x)$, а не разностные операторы $\Delta_\pm$.

Одна из наиболее ранних таких теорем принадлежит Фридрихсу [1954]:

Предположим, что при $\Delta t = \Delta x$ разностная схема

$$u^{n+1} = Cu^n = \left[\sum_{\beta} c^{(\beta)}(x) T^\beta \right] u^n \quad (5.63)$$

такова, что (1) матрицы $c^{(\beta)}(x)$ эрмитовы, не зависят от Δt и удовлетворяют условию Липшица по x и (2) матрицы $c^{(\beta)}(x)$, кроме того, неотрицательны; тогда эта схема устойчива.

Как было отмечено Лаксом [1961], условие (2) весьма стеснительно — ему удовлетворяют только разностные схемы первого порядка, — но зато оно позволяет очень просто доказать теорему. Заметим прежде всего, что из него следует, что норма символа разностной схемы не превосходит единицы:

$$\|G(x, \xi)\| = \left\| \sum_{\beta} c^{(\beta)}(x) e^{i\beta \cdot \xi} \right\| \leq 1 \quad \text{при всех } x \text{ и } \xi. \quad (5.64)$$

Действительно, в силу условия согласованности $\sum_{\beta} c^{(\beta)}(x)$ есть единичная матрица; поэтому если α_{β} — любые комплексные числа, по модулю равные единице, то из предположения (2) вытекает, что

$$\begin{aligned} \left| v^* \sum_{\beta} \alpha_{\beta} c^{(\beta)} u \right| &\leq \sum_{\beta} |v^* c^{(\beta)} u| \leq \frac{1}{2} \sum_{\beta} (v^* c^{(\beta)} v + u^* c^{(\beta)} u) = \\ &= \frac{1}{2} |v|^2 + \frac{1}{2} |u|^2; \end{aligned}$$

полагая $v = \sum_{\beta} d_{\beta} c^{(\beta)} u$, получаем нужный результат $|v|^2 \leq |u|^2$.
Далее,

$$\begin{aligned} |(v, Cu)| &= \left| \sum_{\beta} \int v^*(x) c^{(\beta)}(x) u(x + \beta \Delta t) dx \right| \leq \\ &\leq \frac{1}{2} \|v\|^2 + \frac{1}{2} \sum_{\beta} \int u^*(x) c^{(\beta)}(x - \beta \Delta t) u(x) dx \leq \\ &\leq \frac{1}{2} \|v\|^2 + \frac{1}{2} \sum_{\beta} \int u^*(x) c^{(\beta)}(x) u(x) dx + O(\Delta t) = \\ &= \frac{1}{2} \|v\|^2 + \frac{1}{2} \|u\|^2 + O(\Delta t). \end{aligned}$$

Полагая $v = Cu$, получаем окончательно

$$\|C\| \leq 1 + O(\Delta t). \quad (5.65)$$

Мы уже упоминали первую работу Лакса, посвященную теоремам такого типа. Позднее Лакс и Ниренберг [1966] доказали теорему, устанавливающую прямую связь между локальным условием устойчивости (5.64) и условием устойчивости (5.65) в целом:

Предположим, что действительные симметричные матричные коэффициенты $c^{(\beta)}(x)$ не зависят от Δt и имеют ограниченные вторые производные по x , а условие (5.64) выполнено всюду. Тогда для действительных векторных функций u^n разностная схема (5.63) устойчива.

Первоначальное доказательство годилось лишь для скалярного случая, и было основано на использовании разложения единицы, зависящего от Δt . Доказательство для векторного случая даже более изящно. Сначала введем эрмитову матрицу

$$R(x, \xi) = I - G^*(x, \xi) G(x, \xi) \quad (5.66)$$

и построим соответствующий разностный оператор R_Δ . Из условия Липшица для коэффициентов $c^{(\beta)}$ следует, что

$$R_\Delta = I - C^*C + O(\Delta t). \quad (5.67)$$

Поэтому нам надо, исходя из предположения (5.64), т. е. из того, что

$$R(x, \xi) \geq 0 \text{ для всех } x \text{ и } \xi, \quad (5.68)$$

доказать, что для некоторой константы M

$$(u, R_\Delta u) \geq -M \Delta t \|u\|^2. \quad (5.69)$$

Для этого нам потребуются две леммы.

Лемма 1. Если символ R эрмитов и $\varphi(x)$ — гладкая действительная скалярная функция, для которой $|\varphi(x)| \leq 1$, то для действительных u

$$|(\varphi u, R_\Delta \varphi u) - (\varphi u, \varphi R_\Delta u)| \leq K(L + L^2)(\Delta t)^2 \|u\|^2, \quad (5.70)$$

где постоянная K зависит только от R , а L — постоянная Липшица для φ .

Так как символ R эрмитов, то оператор R_Δ есть сумма членов вида $r(x)[T^\beta \pm T^{-\beta}]$, где матрица r эрмитова в случае положительного знака и косоэрмитова в противном случае. Рассмотрим, к примеру, первый случай. Используя индекс β для обозначения сдвига $x \rightarrow x + \beta \Delta t$, левую часть неравенства (5.70) для рассматриваемого члена можно оценить так:

$$\begin{aligned} & \left| \int (\varphi_\beta - \varphi) \varphi u^T r u_\beta dx - \int (\varphi - \varphi_{-\beta}) \varphi u^T r u_{-\beta} dx \right| = \\ & = \left| \int (\varphi_\beta - \varphi) (\varphi u^T r u_\beta - \varphi_\beta u_\beta^T r_\beta u) dx \right| \leq \\ & \leq L |\beta| \Delta t \int |(\varphi - \varphi_\beta) u^T r u_\beta + \varphi_\beta (u^T r u_\beta - u_\beta^T r_\beta u)| dx, \end{aligned}$$

где u^T — матрица, транспонированная к u . Первый член в подынтегральном выражении в правой части не превосходит $\|r\|L|\beta|\|u\|^2\Delta t$, а второй ограничен произведением постоянной Липшица для $r(x)$ на $|\beta|\|u\|^2\Delta t$, так как $u^T r u_\beta = u_\beta^T r u$; отсюда и следует нужный результат.

Лемма 2. Если матрицы A и B эрмитовы, $A + B \geq 0$ и $A - B \geq 0$, то для любого действительного числа M

$$|(u, Bv)| \leq \frac{1}{2} M^2 (u, Au) + \frac{1}{2} M^{-2} (v, Av). \quad (5.71)$$

Для доказательства достаточно заметить, что

$$\begin{aligned} ((f + g), (A + B)(f + g)) + ((f - g), (A - B)(f - g)) = \\ = 2(f, Af) + 2(g, Ag) + 2(f, Bg) + 2(g, Bf) \geq 0, \end{aligned}$$

положить $f = Mu$, $g = M^{-1}e^{i\theta}v$ и выбрать подходящее θ .

Для доказательства нашей теоремы возьмем теперь некоторое основное разбиение единицы, $\sum_i d_i^2(x) \equiv 1$, и рассмотрим семейство функций $e_i(x) = d_i((\Delta t)^{-1/2}x)$. Функции $e_i(x)$ имеют носитель порядка $(\Delta t)^{1/2}$ и удовлетворяют условию Липшица: $|e_i(x) - e_i(y)| \leq \text{const} (\Delta t)^{-1/2} |x - y|$. Мы утверждаем, что

$$|(u, R_\Delta u) - \sum_i e_i u, R_\Delta e_i u| \leq \text{const} \Delta t \|u\|^2.$$

В самом деле каждый член $r(x)T^\beta$ оператора R_Δ приводит к разности

$$\begin{aligned} \int u^T r u_\beta \left(1 - \sum_i e_i(x) e_i(x + \beta \Delta t)\right) dx = \\ = \int u^T r u_\beta \sum_i \frac{1}{2} [e_i(x) - e_i(x + \beta \Delta t)]^2 dx. \end{aligned}$$

Разность в квадратных скобках в последнем выражении при каждом фиксированном значении x отлична от нуля только для конечного числа значений i , поэтому вся сумма является величиной порядка $O[(\Delta t)^{1/2}]$, откуда и следует справедливость сделанного выше утверждения.

Таким образом, нам осталось получить неравенство (5.69) для функции $v = e_i u$, имеющей носитель порядка $(\Delta t)^{1/2}$. Предполагая, что начало координат принадлежит этому носителю, представим на нем R в виде ряда Тэйлора

$$R(x, \xi) = R(0, \xi) + \sum_{j=1}^d x_j R^{(j)}(\xi) + O(\Delta t), \quad (5.72)$$

где $R^{(j)}(\xi) = \partial_{x_j} R(x, \xi)|_{x=0}$; если теперь положить

$$\begin{aligned} R^{(0)}(\xi) &= R(0, \xi), \\ \tilde{R}(x, \xi) &= R^{(0)}(\xi) + \sum_j x_j R^{(j)}(\xi), \end{aligned}$$

то ясно, что наша задача сводится к тому, чтобы доказать, что

$$(v, \tilde{R}_\Delta v) \geq -M \Delta t \|v\|^2,$$

т. е. что

$$(\mathbf{v}, R_{\Delta}^{(0)} \mathbf{v}) + \sum_j (\mathbf{w}_j, R_{\Delta}^{(j)} \mathbf{v}) \geq -M \Delta t \|\mathbf{v}\|^2,$$

где $\mathbf{w}_j = x_j \mathbf{v}$, а значок Δ употребляется для обозначения разностного оператора, отвечающего соответствующему символу. Далее, так как $R(x, \xi) \geq 0$, то должна существовать такая постоянная M_1 , что

$$R^{(0)}(\xi) + x_j R^{(j)}(\xi) + M_1 \Delta t I \geq 0 \quad \text{при} \quad |x_j| \leq (\Delta t)^{1/2}.$$

Таким образом, мы можем применить лемму 2 к

$$A = R^{(0)}(\xi) + M_1 \Delta t I, \quad B = (\Delta t)^{1/2} R^{(j)}(\xi)$$

и получить оценку «трудных» членов в $R^{(j)}$. Так как теперь все операторы суть операторы с постоянными коэффициентами, то эта оценка непосредственно переносится в исходное вещественное пространство. В результате получаем

$$(\Delta t)^{1/2} |(\mathbf{w}_j, R_{\Delta}^{(j)} \mathbf{v})| \leq \frac{M_2}{2} [(\mathbf{v}, R_{\Delta}^{(0)} \mathbf{v}) + M_1 \Delta t \|\mathbf{v}\|^2] + \\ + \frac{1}{2M_2} [(\mathbf{w}_j, R_{\Delta}^{(0)} \mathbf{w}_j) + M_1 \Delta t \|\mathbf{w}_j\|^2].$$

Если взять $M_2 = (\Delta t)^{1/2}/d$ и заметить, что $\|\mathbf{w}_j\|^2 \leq O(\Delta t) \|\mathbf{v}\|^2$, то мы придем к неравенству

$$|(\mathbf{w}_j, R_{\Delta}^{(j)} \mathbf{v})| \leq \frac{1}{2d} (\mathbf{v}, R_{\Delta}^{(0)} \mathbf{v}) + \frac{d}{2\Delta t} (\mathbf{w}_j, R_{\Delta}^{(0)} \mathbf{w}_j) + O(\Delta t) \|\mathbf{v}\|^2,$$

так что остается доказать, что

$$\frac{1}{2} (\mathbf{v}, R_{\Delta}^{(0)} \mathbf{v}) - \frac{d}{2\Delta t} \sum_j (\mathbf{w}_j, R_{\Delta}^{(0)} \mathbf{w}_j) \geq -O(\Delta t) \|\mathbf{v}\|^2.$$

Используя теперь лемму 1 с $\varphi = x_j$; находим

$$|(\mathbf{w}_j, R_{\Delta}^{(0)} \mathbf{w}_j)| - (x_j^2 \mathbf{v}, R_{\Delta}^{(0)} \mathbf{v}) \leq K (\Delta t)^2 \|\mathbf{v}\|^2,$$

где K зависит только от $R_{\Delta}^{(0)}$; поэтому нужный результат будет следовать из того факта, что

$$([1 - (d/\Delta t) |x|^2] \mathbf{v}, R_{\Delta}^{(0)} \mathbf{v}) \geq -O(\Delta t) \|\mathbf{v}\|^2,$$

если мы докажем, что носитель функции $\mathbf{v}$ содержится в шаре $|x| \leq (\Delta t/2d)^{1/2}$. Чтобы убедиться в последнем, положим $\psi(x) = (1 - (d/\Delta t) |x|^2)^{1/2}$ внутри этого шара и продолжим ψ вне шара как гладкую положительную функцию. Постоянная Лип-

шица для нее равна $(d/\Delta t)^{1/2}$ и в силу той же леммы 1

$$|(\psi^2 v, R_{\Delta}^{(0)} v) - (\psi v, R_{\Delta}^{(0)} \psi v)| \leq 2K (\Delta t)^2 (d/\Delta t) \|v\|^2,$$

в то время как

$$(\psi v, R_{\Delta}^{(0)} \psi v) \geq 0.$$

§ 5.6. Нелинейные уравнения с гладкими решениями

Для нелинейных уравнений пока еще не существует методов исследования достаточно общих классов разностных схем. За исключением отдельных частных случаев, совсем немного доказано относительно разностных схем, аппроксимирующих разрывные решения, которые часто возникают при решении таких уравнений. Только в случае достаточно гладких решений возможен подход с помощью уравнений первой вариации. В гидродинамике, например, это означает, что для аппроксимации ударных волн пока нет строгой теории.

Однако, хотя предположение о гладкости решения дифференциального уравнения и исключает из рассмотрения многие физически интересные случаи, оно все-таки позволяет придать теории, развиваемой в последних трех главах, сколько-нибудь законченный вид. Хотя понятие устойчивости как ограниченности решений разностных уравнений остается в силе, теорема Лакса об эквивалентности становится, конечно, неприменимой, поэтому доказывать сходимость приходится непосредственно. Это можно сделать различными способами. Весьма простой способ рассуждений, изложенный здесь, принадлежит Стрэнгу [1964a].

Рассмотрим гиперболическую краевую задачу для квазилинейной системы

$$\frac{\partial u}{\partial t} = \sum_{j=1}^d A_j(u, x) \frac{\partial u}{\partial x_j}, \quad -\infty < x_i < +\infty, \quad 0 \leq t \leq 1, \quad (5.73)$$

где $u = u(x, t)$ есть p -мерная векторная функция с начальным значением $u(x, 0) = u_0(x)$. Аппроксимирующую эту задачу разностную схему запишем в виде

$$u^{n+1} = \varphi(u^n, x, \Delta t), \quad u^0(x, \Delta t) = u_0(x), \quad (5.74)$$

где правая часть является нелинейной функцией от значений $u^n(x + \beta \Delta t)$ функции u^n в конечном числе точек, соседних с x . Как и в предыдущем параграфе, выберем все пространственные приращения равными Δt . Тогда в силу согласованности для любой гладкой функции v имеем

$$\varphi(v, x, \Delta t) = v + \Delta t \sum_{j=1}^d A_j(v, x) \frac{\partial v}{\partial x_j} + o(\Delta t); \quad (5.75)$$

порядок точности схемы равен q , если для решения уравнения (5.73)

$$\|u(x, t + \Delta t) - \varphi(u, x, \Delta t)\| = O[(\Delta t)^{q+1}]. \quad (5.76)$$

Предположим теперь, что мы подставили разложение

$$U(x, t, \Delta t) = u(x, t) + \sum_m (\Delta t)^m U_m(x, t) \quad (5.77)$$

в равенство $U(x, t + \Delta t, \Delta t) = \varphi(U(x, t, \Delta t), x, \Delta t)$, разложим его левую часть в ряд Тэйлора и приравняем нулю коэффициенты при каждой степени Δt , используя при этом (5.75). Первое из полученных таким образом уравнений в точности совпадает с дифференциальным уравнением (5.73) для функции u , а остальные будут линейными уравнениями относительно U_m :

$$\frac{\partial U_m}{\partial t} = \sum_{j=1}^d A_j(u, x) \frac{\partial U_m}{\partial x_j} + \alpha(x, t) U_m + \beta_m(x, t), \quad (5.78)$$

$$U_m(x, 0) = 0, \quad m = 1, 2, \dots$$

Здесь коэффициент α зависит от первых производных функции u и первых производных по u от функций A_j , а β_m зависит от производных до порядка $m + 1$ от u , φ и A_j и до порядка $m + 1 - j$ от U_j для $j < m$. Далее, вследствие гиперболичности уравнения (5.73) решения уравнений (5.78) существуют и непрерывно зависят от неоднородных членов β_m и их производных до некоторого порядка r_0 . Поэтому в предположении, что u , φ и A_j имеют $r_0 + r + 1$ производных, все U_m , $m = 1, 2, \dots, r$, существуют; они называются *главными членами ошибки*¹⁾ и функция

$$U(x, t, \Delta t) = u(x, t) + \sum_{m=1}^r (\Delta t)^m U_m(x, t) \quad (5.79)$$

удовлетворяет соотношению

$$U(x, t + \Delta t, \Delta t) = \varphi(U(x, t, \Delta t), x, \Delta t) + o[(\Delta t)^{r+1}]. \quad (5.80)$$

Основная идея Стрэнга состоит в том, чтобы сравнивать с разностной аппроксимацией u^n функцию U вместо u . Преимущество этого подхода в том, что за счет выбора достаточно большого r возмущение разностного уравнения (5.74) можно сделать сколь угодно малым. Это нужно потому, что при оценке разности между φ и ее первой вариацией нам придется переходить от среднеквадратичной нормы к норме максимум модуля.

¹⁾ Заметим, что эти члены ошибки не зависят от Δt ; следовательно, то же верно и для оценок их производных, которые потребуются ниже.

Итак, обозначим через E^n разность $U(n\Delta t) - u^n$ и вычтем уравнение (5.74) из (5.80) при $t = n\Delta t$; для каждого β обозначим через $c^\beta(v)$ якобиан функции $\varphi(v, x, \Delta t)$ относительно соответствующих переменных, отвечающих $v(x + \beta\Delta t)$ и, используя теорему о среднем значении, получим

$$E^{n+1}(x) = \sum_{\beta} c^\beta(U(n\Delta t) - \theta E^n) E^n(x + \beta\Delta t) + o[(\Delta t)^{r+1}], \quad (5.81)$$

$$E^0(x) = 0,$$

где θ лежит между 0 и 1 и зависит от n , β и x . Линейный разностный оператор $C^n = C(n\Delta t, \Delta t)$, соответствующий коэффициентам c^β , вычисленным в точке $u(x, n\Delta t)$, называется *первой вариацией* оператора φ . Предположим, что он устойчив, т. е. что

$$\left\| \prod_{i=1}^n C^i \right\| \leq K \quad \text{при} \quad n\Delta t \leq 1.$$

Докажем по индукции, что если порядок точности оператора φ равен q и r взято равным $q + [(d+1)/2]$, где $[s]$ обозначает целую часть действительного числа s , то

$$\max_x |E^n(x, \Delta t)| \leq n(\Delta t)^{q+1} \quad \text{при} \quad n\Delta t \leq 1. \quad (5.82)$$

Очевидно, это верно для $n = 0$; предположим, что это справедливо для $n < N$. Перепишем (5.81) в виде

$$E^{n+1} = (C^n + B^n) E^n + o[(\Delta t)^{r+1}].$$

Из (5.79) и того факта, что $q \geq 1$, следует, что для всех $n < N$

$$\|B^n\| \leq \mu \Delta t, \quad (5.83)$$

где μ — некоторая константа. Следовательно, вспоминая рассуждения из § 3.9, получаем

$$E^N = \sum_{i=1}^{N-1} \prod_{j=i}^{N-1} (C^j + B^j) o[(\Delta t)^{r+1}]$$

и

$$\left\| \prod_{j=i}^{N-1} (C^j + B^j) \right\| \leq K e^{\mu K}. \quad (5.84)$$

Это произведение при $i = 1$ представляет собой матрицу бесконечного порядка, у которой в каждой строке имеется $O(N^d)$ отличных от нуля элементов, «связывающих» значение E^N в некоторой точке N -го временного слоя с ошибками аппроксимации в $O(N^d)$ точках первого временного слоя. Это — оценка в L_2 -норме. Оценка соответствующей нормы максимум модуля получает-

ся умножением на $O(N^{d/2})$, так что

$$\max_x |E^N| \leq N K e^{\mu K} O(N^{d/2}) o[(\Delta t)^{r+1}] \leq N (\Delta t)^{q+1}$$

при $\Delta t < \delta$, где δ не зависит от N . Этим завершаются индукция и доказательство сходимости $u^n(x)$ к $u(x, n\Delta t)$ при $\Delta t \rightarrow 0$. Сформулируем результат в виде теоремы.

Предположим, что оператор Φ из уравнения (5.74) является согласованным и его первая вариация устойчива в L_2 . Тогда если u , A_j и Φ имеют непрерывные производные до порядка $[(d+1)/2] + r_0 + 2$, то при $\Delta t \rightarrow 0$ разностная аппроксимация u^n сходится к решению уравнения (5.73).

Рассматривая порядок величины членов β_m , можно показать, что если Φ имеет порядок точности q и предполагается существование $[(d+1)/2] + r_0 + q + 1$ производных, то $u^n(x) - u(x, n\Delta t) = O[(\Delta t)^q]$.

Несмотря на общность этой теоремы, мы должны предостеречь читателя от мысли, что относительно нелинейных уравнений с достаточно гладкими решениями все ясно. О том, что это не так, свидетельствуют серьезные ограничения, с которыми приходится сталкиваться при применении идей устойчивости и сходимости в практических вычислениях. Одна из причин здесь та, что влияние граничных условий пока никак не учитывалось. Следующий пример, принадлежащий Рихтмайеру [1963], показывает, какого рода неприятности могут возникнуть даже в самых простых случаях.

Дифференциальное уравнение

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left(\frac{1}{2} u^2 \right) = 0, \quad 0 \leq x \leq 1, \quad t \geq 0, \quad (5.85)$$

с начальными и граничными условиями

$$\begin{cases} u(x, 0) = x, & 0 \leq x \leq 1, \\ u(0, t) = 0, & t \geq 0, \end{cases} \quad (5.86)$$

имеет точное решение

$$u(x, t) = x/(1+t). \quad (5.87)$$

Предположим, что мы пытаемся решить эту задачу, используя «петлеобразное» разностное уравнение

$$u_j^{n+1} - u_j^{n-1} + \frac{\lambda}{2} [(u_{j+1}^n)^2 - (u_{j-1}^n)^2] = 0 \quad (5.88)$$

при $\Delta t = \lambda \Delta x$. Можно показать, что это уравнение имеет решения, которые для фиксированного t возрастают как $\exp(\text{const}/\Delta t)$ при $\Delta t \rightarrow 0$, даже если для линеаризованного уравнения выпол-

нено условие фон Неймана

$$\lambda \max_{x, t} \{|u|\} < 1. \quad (5.89)$$

Действительно, рассмотрим решение вида

$$u_j^n = C^n \cos \frac{\pi j}{2} + S^n \sin \frac{\pi j}{2} + U^n \cos \pi j + V. \quad (5.90)$$

Подстановка его в уравнение приводит к рекуррентным соотношениям

$$\begin{aligned} C^{n+1} - C^{n-1} &= 2\lambda S^n (U^n - V), \\ S^{n+1} - S^{n-1} &= 2\lambda C^n (U^n + V), \\ U^{n+1} - U^{n-1} &= 0. \end{aligned} \quad (5.91)$$

Согласно третьему из них U^n принимает два циклически повторяющихся постоянных значения, скажем A и B . Поэтому первые два соотношения дают уравнение

$$C^{n+2} - 2C^n - C^{n-2} = 4\lambda^2 (A + V)(B - V) C^n \quad (5.92)$$

и аналогичное уравнение для S^n . Если коэффициент в правой части заключен между -4 и 0 , то все решения будут ограничены; в противном случае всегда найдется такое решение C^n , что $|C^n|$ будет расти как экспонента от n , т. е. от $t/\Delta t$.

Чтобы увидеть, когда может появиться такое экспоненциально возрастающее решение, заметим, что в силу условия фон Неймана (5.89) справедливо неравенство

$$\lambda \{|V| - \max(|A|, |B|)\} < 1,$$

которое показывает, что коэффициент при C^n в правой части (5.92) не может быть меньше -4 . Если кроме того

$$|A| < |V| \text{ и } |B| < |V|, \quad (5.93)$$

то этот коэффициент не может быть больше нуля. Поэтому если мы будем рассматривать постоянный член V в (5.90) как гладкое решение, для которого другие члены являются возмущением, и если амплитуда этого возмущения не превосходит предельного значения, или порога, указанного в (5.93), то оно не будет возрастать; но если первоначальное возмущение больше этого порога, то можно ожидать его экспоненциального роста.

Контрольные расчеты, проведенные для задачи (5.85), (5.86) Рихтмайером и Мортонем [1964], а также Стеттером [1965], подтверждают наличие этого и ряда других сходных явлений. Использование петлеобразной схемы (5.88) в некоторых случаях приводит к решениям, которые сильно возрастают, в то время как схема для соответствующей линеаризованной задачи

оказывается либо устойчивой, либо приводит лишь к очень незначительному возрастанию ошибки. Выводы из этих примеров не совсем ясны, отчасти из-за влияния граничных условий. Возможное объяснение состоит в том, что в обоих случаях петлеобразная схема обеспечивает сходимость аппроксимации *при достаточно малых Δt* , т. е. тогда, когда ошибка аппроксимации в разностной схеме порождает достаточно малые возмущения. Но для нелинейной задачи поведение решения существенно зависит от относительной величины возмущений, что не имеет места для линейной задачи.

Во всяком случае практическое различие между этими задачами весьма велико. Из последующих глав будет видно, что во многих случаях наличия устойчивости при $\Delta t \rightarrow 0$ отнюдь не достаточно для практических целей.

Следует отметить, что даже после линеаризации петлеобразная схема находится только «на грани» устойчивости, так как собственные значения ее матрицы перехода по модулю равны единице. Поэтому к ней невозможно применить изложенные выше теоремы Крайса. Однако в (5.88) можно ввести дополнительные разностные члены таким образом, что схема станет диссипативной в указанном выше смысле без изменения порядка точности. Рихтмайер и Мортон показали, что коэффициенты этих членов можно выбрать так, чтобы обеспечить устойчивость даже для нелинейной задачи. Нужно также отметить, что хотя здесь рассматривается только случай скалярного переменного, но теорема Лакса и Ниренберга неприменима даже к линеаризованной системе. Петлеобразная схема двухслойна, и если ее привести к однослойной схеме (см. гл. 7), то вектор решения будет иметь две компоненты; матрица перехода для такой системы не удовлетворяет условию теоремы Лакса и Ниренберга.

Глава 6

СМЕШАННЫЕ КРАЕВЫЕ ЗАДАЧИ

§ 6.1. Введение

Специалист по численному анализу, перед которым поставлена проблема выбора устойчивых конечноразностных схем, может идти двумя существенно различными путями: попытаться получить такой критерий, который был бы пригоден для определения устойчивости любой из предложенных ему разностных схем, или же использовать те методы, устойчивость которых уже установлена. Первый путь более удовлетворяет честолюбию математика, и до сих пор при исследовании устойчивости мы шли именно по этому пути. Однако из последней главы видно, что встречающиеся при этом трудности непрерывно возрастают. Так, для линейных гиперболических уравнений в теоремах общего характера надо предполагать либо большую гладкость коэффициентов плюс выполнение локального условия устойчивости, либо наличие диссипативности. В обоих случаях доказательства теорем достаточно сложны. По-видимому при фактическом исследовании линеаризованных задач для нелинейных уравнений условие диссипативности является более подходящим, поскольку численные эксперименты, типа упомянутых в § 5.6, наводят на мысль, что оно весьма желательно, даже если и не является строго необходимым. Но для многих имеющих практическое значение линейных разностных схем условия этих общих теорем оказываются слишком стеснительными: действительно, во многих случаях было бы удивительно, если бы условие Липшица для коэффициентов давало в смысле устойчивости что-либо другое по сравнению с случаем постоянных коэффициентов. Таким образом, как это часто бывает, выигрывая в общности, мы проигрываем в точности для конкретных частных случаев.

Поэтому в этой главе мы уделим больше внимания изучению конкретных типов разностных схем. На этом пути мы сможем кроме того начать рассмотрение задач с непериодическими граничными условиями для ограниченных областей и изучить влияние этих граничных условий на устойчивость. Введение границ помимо необходимости использовать дополнительный аналитический аппарат приводит еще к двум новым затруднениям при исследовании устойчивости. Во-первых, для того что-

бы показать, что для разностного уравнения не существует собственных колебаний, возрастающих от шага к шагу, нужно теперь рассматривать колебания, отличные от колебаний с постоянной амплитудой, т. е. от компонент Фурье для задач с постоянными коэффициентами; действительно, среди колебаний, которые убывают при удалении от границы, также могут быть собственные колебания, разумеется, если они удовлетворяют граничным условиям. В действительности этими последними условиями исключаются все колебания, возрастающие со временем, в предположении, что устойчивость (или корректная постановка задачи) обеспечена. Во-вторых, многие разностные схемы требуют для своей реализации большего числа граничных условий, чем это необходимо при решении соответствующей задачи для дифференциального уравнения. Поэтому возникает задача научиться выбирать эти дополнительные условия таким образом, чтобы при этом не ухудшить устойчивость.

§ 6.2. Основные идеи энергетического метода

Начнем рассмотрение наших задач с несколько более подробного, чем до сих пор, изучения так называемого энергетического метода. Это несколько странное название относится к широкому кругу технических приемов, основанных на понятии сильной устойчивости, о которой говорилось в конце § 5.2. Основная идея состоит в том, чтобы построить для вектора решения такую норму, в которой он возрастал бы за каждый шаг на множитель, не превосходящий $1 + O(\Delta t)$. Это означает устойчивость в такой норме, и все дальнейшие рассуждения обычно сводятся к тому, чтобы установить эквивалентность этой нормы и L_2 -нормы. Причиной указанного названия является то, что в некоторых наиболее простых случаях в качестве такой нормы можно взять физическую энергию системы.

Подобные методы уже давно применяются в теории дифференциальных уравнений для доказательства существования и единственности решений (см., например, работы Фридрихса [1954, 1958] и Р. Филлипса [1957, 1959], в которых имеются ссылки на более ранние публикации). Их применение к разностным уравнениям также имеет долгую историю, начинающуюся с работы Куранта, Фридрихса и Леви [1928]. Затем они использовались многими авторами, в том числе Фридрихсом [1954], Лисом [1960], Лаксом [1961] и Крайсом [1963].

Иногда с помощью этих методов можно получать совершенно общие результаты; мы уже делали это в предыдущей главе. В начале своих исследований по задачам с переменными коэффициентами Лакс [1960] рассмотрел вопрос о границах применимости энергетического метода и показал, что при определен-

ных условиях норма с нужными свойствами всегда существует и может быть выражена в явном виде через решение и определенное число разностей от него. Однако условия, при которых это возможно, довольно ограничительны. Из теоремы Крайса о матрицах мы знаем, что такие нормы существуют по крайней мере для задач с постоянными коэффициентами; трудность заключается в их точном представлении с помощью решения и разностей от него.

Поэтому мы будем рассматривать энергетический метод скорее как совокупность приемов, применяемых к каждой разностной схеме (или классу разностных схем) индивидуально для получения нормы с нужными свойствами для этой конкретной схемы. Так, для одного уравнения $B_1 u^{n+1} = B_0 u^n$ с одной пространственной переменной весь процесс состоит по существу из двух этапов. Сначала, возводя в квадрат обе части уравнения и суммируя по всем точкам сетки, т. е. выбирая подходящее скалярное произведение, мы стремимся получить соотношение вида

$$S_{n+1} = S_n + O(\Delta t) (\|u^n\|^2 + \|u^{n+1}\|^2), \quad (6.1)$$

где $S_n = S(u^n)$ — некоторая квадратичная форма от решения u^n и его разностей. Затем мы должны показать, что форма S_n является положительно определенной и эквивалентна квадрату L_2 -нормы u^n ; именно здесь в качестве условия устойчивости появляется некоторое ограничение на шаг по времени. Заметим, что при рассмотрении разности между u^n и решением дифференциального уравнения можно с помощью этого метода непосредственно получить оценку роста ошибки и скорость сходимости.

Ясно, что получаемое таким образом условие устойчивости в общем случае будет только достаточным. Можно не получить действительных границ устойчивости метода либо вследствие недостаточной изобретательности при построении нормы, либо вследствие ограничений, присущих самому энергетическому методу. Таким образом, этот метод до некоторой степени дополняет анализ на собственные значения, который обычно дает необходимые условия устойчивости. Компенсацией за эти ограничения служит главным образом та простота, с которой можно учесть переменность коэффициентов и граничные условия. Помимо этого, ниже будут построены примеры, показывающие, что понятие устойчивости, определенной до сих пор как свойство, проявляющееся в пределе при $\Delta t \rightarrow 0$, может быть причиной недоразумений на практике. Энергетический метод позволяет ответить на весьма важный практический вопрос — для каких значений Δt при данном конечном Δx норма решения возрастает не быстрее экспоненты от времени t ?

В соответствии с общим принципом тщательной подгонки нормы к каждому рассматриваемому методу указанные положения лучше всего проверить на конкретных примерах, часть которых приводится в следующем параграфе. Эти примеры иллюстрируют также еще одну особенность энергетического метода, а именно что разностную схему можно (и часто нужно) выбирать так, чтобы облегчить применение метода. В особенности это относится к многомерным задачам.

Прежде всего введем некоторые обозначения и получим несколько основных неравенств и тождеств. Для простоты ограничимся здесь задачами лишь с одним пространственным переменным x , для которого возьмем в качестве стандартной области отрезок $0 \leq x \leq 1$; если одна или обе границы находятся в бесконечности, то обычно это приводит только к упрощениям, и необходимые при этом изменения в приводимых ниже утверждениях очевидны. Покроем область изменения x равномерной сеткой с шагом Δx , положение которой относительно границ выбрано и фиксировано, и будем считать, что решение $u^n(x)$ ищется только в узлах сетки, где оно обозначается через u_j^n ; j пробегает значения от 0 до J . Выбор сетки в крайних точках в значительной степени определяется заданными граничными условиями, и мы рассмотрим здесь только две возможности¹⁾: либо крайняя точка является точкой сетки, называемой в этом случае *граничной точкой*, что обычно бывает тогда, когда в качестве граничного условия задано само значение функции; либо она лежит посередине между двумя точками сетки, что чаще всего бывает тогда, когда в качестве граничного условия задается производная от неизвестной функции. Точки сетки внутри области называются *внутренними точками*. Во втором случае иногда удобно ввести фиктивные точки вне области изменения x , называемые *внешними точками*. Отметим, что значения, которые может принимать Δx , ограничены выбором сетки на границах: например, если обе крайние точки являются граничными, то должно быть $(\Delta x)^{-1} = J$.

В литературе существует большой разноречивостью в обозначениях для конечных разностей. Для наших целей удобнее обозначать разности с помощью явного оператора, а не с помощью, скажем, дополнительного нижнего индекса у функции, от которой берется разность. Кроме того, мы предпочитаем использовать для неразделенных разностей обозначения, близкие к тем, которые применялись в старых руководствах по данному пред-

¹⁾ В практических задачах для различных компонент вектора u_j^n сетки могут значительно отличаться друг от друга; их выбор зависит от удобства аппроксимации дифференциальных уравнений и граничных условий; здесь подобные усложнения не будут рассматриваться.

мету. Хотя большая часть обозначений уже была введена в различных местах книги, стоит собрать их все вместе:

$$\Delta_+ u_j = u_{j+1} - u_j, \quad \Delta_- u_j = u_j - u_{j-1}, \quad (6.2)$$

$$\Delta_0 u_j = \frac{1}{2} (\Delta_+ + \Delta_-) u_j = \frac{1}{2} (u_{j+1} - u_{j-1}), \quad (6.3)$$

$$\delta^2 u_j = \Delta_+ \Delta_- u_j = \Delta_- \Delta_+ u_j = u_{j+1} - 2u_j + u_{j-1}. \quad (6.4)$$

Если имеется несколько пространственных переменных, то они различаются введением у разностных операторов нижних индексов. Кроме того, мы иногда будем использовать операторы сдвига

$$T_+ u_j = u_{j+1}, \quad T_- u_j = u_{j-1}, \quad (6.5)$$

как, например, в тождествах для разностей от произведения:

$$\Delta_+ (au) = a\Delta_+ u + (\Delta_+ a) T_+ u = T_+ (a\Delta_- u) + (\Delta_+ a) u, \quad (6.6)$$

$$\Delta_- (au) = a\Delta_- u + (\Delta_- a) T_- u = T_- (a\Delta_+ u) + (\Delta_- a) u, \quad (6.7)$$

$$\Delta_0 (au) = a\Delta_0 u + (\Delta_0 a) u + \frac{1}{2} \Delta_- [(\Delta_+ a) (\Delta_+ u)]. \quad (6.8)$$

Хотя эти тождества приведены лишь в скалярной форме, ясно, что они справедливы также и в том случае, когда a является матрицей, а u — вектором.

Для функции u^n , не определенной в каждой точке x -пространства, мы не можем использовать обычное интегральное определение L_2 -нормы. Вместо этого возьмем в качестве нормы следующую сумму:

$$(u, v)_\Delta = \Delta x \sum_j u_j^* v_j, \quad \|u\|_\Delta^2 = (u, u)_\Delta, \quad (6.9)$$

где суммирование распространяется только на внутренние точки сетки; если нет опасности спутать эту норму с обычной L_2 -нормой, то индекс Δ опускается. Может показаться, что это не согласуется с принципом гл. 2 вложения сеточных функций u^n в функциональное пространство дифференциальной задачи. Но в действительности это не так. Например, в случае когда у сетки нет граничных точек, т. е. граница находится в бесконечности или проходит посередине между двумя точками сетки, то в качестве результата такого вложения функции u^n можно взять ступенчатую функцию со значением u_j^n на каждой половине отрезка сетки по обе стороны от ее j -го узла; тогда норма (6.9) совпадает с L_2 -нормой. С другой стороны, если точки $x = 0$ и $x = 1$ являются граничными и в каждой из них задано однородное условие $u^n = 0$, то формула (6.9) эквивалентна аппроксимации соответствующих интегралов по правилу трапеции. Но если функция u^n продолжена на весь отрезок с помощью линейной интерполяции, то вклад j -го отрезка

в интеграл, определяющий L_2 -норму, равен

$$\int_0^1 \| \mathbf{u}_j^n + (\mathbf{u}_{j+1}^n - \mathbf{u}_j^n) t \|^2 dt = 1/3 (\| \mathbf{u}_j^n \|^2 + \| \mathbf{u}_{j+1}^n \|^2 + \operatorname{Re} \mathbf{u}_j^{n*} \mathbf{u}_{j+1}^n)$$

и заключен между одной шестой и половиной от $\| \mathbf{u}_j^n \|^2 + \| \mathbf{u}_{j+1}^n \|^2$. Следовательно, в этом случае обе нормы также эквивалентны. Можно показать, что они эквивалентны и в промежуточных случаях. Все, что надо делать, — это видоизменять способ вложения в соответствии с рассматриваемой задачей таким образом, чтобы норма (6.9) была эквивалентна L_2 -норме.

В случае когда в скалярных произведениях встречаются разности от $\mathbf{u}$, может возникнуть надобность в изменении пределов суммирования, поэтому удобно ввести обозначения

$$\begin{aligned} (\mathbf{u}, \mathbf{v})_{r,s} &= \Delta x \sum_{j=r}^s \mathbf{u}_j^* \mathbf{v}_j, \\ \| \mathbf{u} \|_{r,s}^2 &= (\mathbf{u}, \mathbf{u})_{r,s}. \end{aligned} \quad (6.10)$$

Для примера легко доказать, что

$$\begin{aligned} \| \Delta_+^n \mathbf{u} \|_{r,s}^2 &\leq 4^n \| \mathbf{u} \|_{r,s+n}^2, \\ \| \Delta_-^n \mathbf{u} \|_{r,s}^2 &\leq 4^n \| \mathbf{u} \|_{r-n,s}^2, \end{aligned} \quad (6.11)$$

тогда как

$$\| \Delta_0^n \mathbf{u} \|_{r,s}^2 \leq \| \mathbf{u} \|_{r-n,s+n}^2. \quad (6.12)$$

Далее если A — матрица, то в силу (6.6) и (6.7) справедливы следующие неравенства:

$$|(\mathbf{u}, \Delta_+(A\mathbf{v}))_{r,s} - (\mathbf{u}, A\Delta_+\mathbf{v})_{r,s}| \leq M_+ \Delta x \| \mathbf{u} \|_{r,s} \| \mathbf{v} \|_{r+1,s+1}, \quad (6.13)$$

$$|(\mathbf{u}, \Delta_-(A\mathbf{v}))_{r,s} - (\mathbf{u}, A\Delta_-\mathbf{v})_{r,s}| \leq M_- \Delta x \| \mathbf{u} \|_{r,s} \| \mathbf{v} \|_{r-1,s-1}, \quad (6.14)$$

$$|(\mathbf{u}, \Delta_0(A\mathbf{v}))_{r,s} - (\mathbf{u}, A\Delta_0\mathbf{v})_{r,s}| \leq M_0 \Delta x \| \mathbf{u} \|_{r,s} \| \mathbf{v} \|_{r-1,s+1}, \quad (6.15)$$

где M с соответствующим индексом есть наибольшее значение нормы матрицы $\partial A / \partial x$ на отрезке, содержащем все точки сетки, входящие в рассматриваемую формулу.

Однако наиболее важными для применений энергетического метода являются формулы суммирования по частям, соответствующие формулам интегрирования по частям в задаче для дифференциального уравнения. Мы уже широко применяли их в тех случаях, когда можно было пренебрегать граничными членами, но теперь мы должны быть более точными. Удобно ввести еще одно обозначение

$$\mathbf{u}_i^* \mathbf{v}_j |_r^s = \mathbf{u}_{s+i}^* \mathbf{v}_{s+j} - \mathbf{u}_{r-i}^* \mathbf{v}_{r-j}. \quad (6.16)$$

Эти формулы могут иметь несколько разновидностей, каждая из которых получается просто суммированием одного из тождеств (6.6) — (6.8); основными являются следующие две формулы:

$$(u, \Delta_+ v)_{r,s} = -(\Delta_- u, v)_{r+1,s+1} + \Delta x \cdot u_0^* v_0 |_r^{s+1}, \quad (6.17)$$

$$(u, \Delta_- v)_{r,s} = -(\Delta_+ u, v)_{r-1,s-1} + \Delta x \cdot u_0^* v_0 |_{r-1}^s; \quad (6.18)$$

объединяя их вместе, получаем

$$(u, \Delta_0 v)_{r,s} = -(\Delta_0 u, v)_{r,s} + \frac{1}{2} \Delta x (u_1^*, v_0 + u_0^* v_1) |_r^s. \quad (6.19)$$

После несложных преобразований приходим к соотношениям

$$(u, \Delta_+ v)_{r,s} = -(\Delta_+ u, v)_{r,s} - (\Delta_+ u, \Delta_+ v)_{r,s} + \Delta x \cdot u_0^* v_0 |_r^{s+1} \quad (6.20)$$

$$(u, \Delta_- v)_{r,s} = -(\Delta_- u, v)_{r,s} + (\Delta_- u, \Delta_- v)_{r,s} + \Delta x \cdot u_0^* v_0 |_{r-1}^s. \quad (6.21)$$

§ 6.3. Простейшие примеры применения энергетического метода: выбор устойчивых аппроксимаций граничных условий и нелинейных членов

В качестве первого примера рассмотрим смешанную краевую задачу для действительной функции $u(x, t)$

$$\frac{\partial u}{\partial t} + a(x) \frac{\partial u}{\partial x} = 0, \quad 0 \leq x < \infty, \quad 0 \leq t < 1, \quad (6.22)$$

с начальными и граничными условиями

$$u(x, 0) = f(x), \quad 0 \leq x < \infty,$$

$$u(0, t) = 0, \quad 0 \leq t \leq 1,$$

где $A \geq a(x) \geq \alpha > 0$. Будем аппроксимировать это уравнение четырехточечной разностной схемой (ср. Томи [1962])

$$\begin{aligned} u_j^{n+1} + u_{j-1}^{n+1} + a_{j-1/2} \Delta_- u_j^{n+1} = \\ = u_j^n + u_{j-1}^n - a_{j-1/2} \Delta_- u_j^n, \quad n = 0, 1, 2, \dots, j = 1, 2, \dots, \end{aligned} \quad (6.23)$$

где положено $\Delta x = \Delta t$, что не уменьшает общности. Чтобы применить энергетический метод, возведем обе части в квадрат и просуммируем по всем j . Полагая

$$\begin{aligned} S_n &= \sum_{j=1}^{\infty} [(u_j^n + u_{j-1}^n)^2 + a_{j-1/2}^2 (u_j^n - u_{j-1}^n)^2] = \\ &= \|(1 + T_-) u^n\|^2 + \|a_{-1/2} \Delta_- u^n\|^2, \end{aligned} \quad (6.24)$$

где $a_{-1/2}$ используется для обозначения $(T_-)^{1/2} a$, получаем

$$\begin{aligned} S_{n+1} - S_n &= -2([1 + T_-] u^n, a_{-1/2} \Delta_- u^n) - \\ &\quad - 2([1 + T_-] u^{n+1}, a_{-1/2} \Delta_- u^{n+1}), \end{aligned} \quad (6.25)$$

Ясно, что

$$(\Delta_- w_j) [1 + T_-] w_j = w_j^2 - w_{j-1}^2 = \Delta_- w_j^2;$$

поэтому

$$S_{n+1} - S_n = -2 (\Delta_- [(u^n)^2 + (u^{n+1})^2], a_{-1/2}).$$

Отсюда после суммирования по частям, используя формулу (6.18) и граничные условия, находим

$$S_{n+1} - S_n = 2 ((u^n)^2 + (u^{n+1})^2, \Delta_+ a_{-1/2}) \leq 2L \Delta x (\|u^n\|^2 + \|u^{n+1}\|^2), \quad (6.26)$$

где L — постоянная Липшица для коэффициента $a(x)$. Следовательно, при постоянном a величина S_n также постоянна, а с другой стороны, она возрастает не более, чем экспоненциально, — во всяком случае это будет так, если мы сможем показать, что определяемая с помощью S_n норма эквивалентна L_2 -норме, но это действительно так, ибо в силу (6.24)

$$4 \min(1, \alpha^2) \|u^n\|^2 \leq S_n \leq 4 \max(1, A^2) \|u^n\|^2. \quad (6.27)$$

Отсюда следует, что

$$\begin{aligned} 4 \min(1, \alpha^2) \|u^n\|^2 \leq S_n \leq S_{n-1} + 2L \Delta x (\|u^n\|^2 + \|u^{n-1}\|^2) \leq \dots \\ \dots \leq S_0 + 4L \Delta x \sum_{m=0}^n \|u^m\|^2, \end{aligned}$$

т. е.

$$\|u^n\|^2 \leq K_1 \|u^0\|^2 + K_2 \Delta x \sum_{m=0}^n \|u^m\|^2,$$

где $K_1 = \max(1, A^2) / \min(1, \alpha^2)$ и $K_2 = L / \min(1, \alpha^2)$. Поэтому

$$\|u^n\|^2 \leq (K_1 + K_2 \Delta x) (1 - K_2 \Delta x)^{-n} \|u^0\|^2 \leq \text{const} \|u^0\|^2$$

при $n\Delta x = n\Delta t \leq 1$ и достаточно малых Δt .

Таким образом, нам удалось доказать безусловную устойчивость разностной схемы, у которой имеется переменный коэффициент и которая не является диссипативной в смысле данного в предыдущей главе определения. Более того, это было сделано для области, ограниченной слева, и доказательство легко можно распространить на случай конечного интервала $0 \leq x \leq 1$, для которого не нужно ставить граничных условий на правом конце ни для дифференциального, ни для разностного уравнений. Если $J\Delta x = 1$, то суммирование при вычислении S_n в (6.24) ведется по $j = 1, 2, \dots, J$, и такие же пределы суммирования берутся для нормы. Неравенства (6.27) не изменяются, за исключением того, что нижняя граница умень-

шается вдвое, так что

$$S_n \geq 2 \min(1, \alpha^2) \left[\sum_{j=1}^J |u_j^n|^2 + \sum_{j=1}^{J-1} |u_j^n|^2 \right] \geq 2 \min(1, \alpha^2) \|u^n\|^2;$$

далее, неравенство (6.26) сохраняется, потому что a не меняет знака¹⁾:

$$\begin{aligned} S_{n+1} - S_n &= -2(\Delta_- [(u^n)^2 + (u^{n+1})^2], a_{-1/2})_{1,J} = \\ &= 2((u^n)^2 + (u^{n+1})^2, \Delta_+ a_{-1/2})_{0,J-1} - 2\Delta x [(u_J^n)^2 + (u_J^{n+1})^2] a_{J-1/2} \leq \\ &\leq 2L\Delta x (\|u^n\|^2 + \|u^{n+1}\|^2). \end{aligned}$$

Последующие рассуждения проходят без всяких изменений. Для удобства ссылок стоит сформулировать эту заключительную часть доказательства в виде отдельной леммы:

Если существуют последовательность действительных чисел S_n и две положительные постоянные K_1 и K_2 такие, что

$$K_1^{-1} \|u^n\|^2 \leq S_n \leq K_1 \|u^n\|^2 \quad (6.28)$$

и

$$S_{n+1} - S_n \leq K_2 \Delta t (\|u^n\|^2 + \|u^{n+1}\|^2), \quad n = 0, 1, 2, \dots, \quad (6.29)$$

то

$$\begin{aligned} \|u^n\|^2 &\leq K_1 (K_1 + 2K_2 \Delta t) (1 - 2K_1 K_2 \Delta t)^{-n} \|u^0\|^2 \leq \\ &\leq \text{const} \|u^0\|^2 \end{aligned}$$

при $n\Delta t \leq T$ и достаточно малых Δt .

В качестве второго примера применим к уравнению (6.22) петлеобразную схему

$$u_j^{n+1} - u_j^{n-1} = -2a_j \Delta_0 u_j^n, \quad j = 1, 2, \dots \quad (6.30)$$

Это трехслойный или двухшаговый метод, общую теорию которого мы отложим до следующей главы, а пока рассмотрим этот простой пример ad hoc. Должны быть заданы начальные значения u^0 и u^1 ; умножив (6.30) на $u_j^{n+1} + u_j^{n-1}$ и просуммировав по $j = 1, 2, \dots$, получим для $n \geq 1$

$$\|u^{n+1}\|^2 - \|u^{n-1}\|^2 = -2(u^{n+1} + u^{n-1}, a \Delta_0 u^n).$$

Положив

$$S_n = \|u^n\|^2 + \|u^{n+1}\|^2 + 2(u^{n+1}, a \Delta_0 u^n), \quad (6.31)$$

получим

$$S_n - S_{n-1} = -2(u^{n-1}, a \Delta_0 u^n) - 2(u^n a \Delta_0 u^{n-1}).$$

¹⁾ Отметим, что это отвечает физической интуиции. Если бы мы имели $a(1) < 0$, то ясно, что нужно было бы ввести граничное условие $u_J^n = 0$.

В силу (6.6), (6.7) и (6.19)

$$\begin{aligned} 2(u^{n-1}, a \Delta_0 u^n) &= 2(u^{n-1}, \Delta_0 (a u^n)) - (u^{n-1}, (T_+ u^n) \Delta_+ a + (T_- u^n) \Delta_- a) = \\ &= -2(\Delta_0 u^{n-1}, a u^n) - (u^{n-1}, (T_+ u^n) \Delta_+ a + (T_- u^n) \Delta_- a). \end{aligned}$$

Следовательно,

$$\begin{aligned} S_n - S_{n-1} &= (u^{n-1}, (T_+ u^n) \Delta_+ a (T_- u^n) \Delta_- a) \leq \\ &\leq 2L \Delta x \|u^{n-1}\| \|u^n\| \leq \\ &\leq L \Delta x (\|u^{n-1}\|^2 + \|u^n\|^2). \end{aligned} \quad (6.32)$$

Далее,

$$2|(u^{n+1}, a \Delta_0 u^n)| \leq A (\|u^{n+1}\|^2 + \|u^n\|^2); \quad (6.33)$$

поэтому если $A < 1$ (в действительности это не что иное, как условие Куранта, Фридрихса и Леви $A \Delta t / \Delta x < 1$), то мы имеем

$$K^{-1} (\|u^n\|^2 + \|u^{n+1}\|^2) \leq S_n \leq K (\|u^n\|^2 + \|u^{n+1}\|^2) \quad (6.34)$$

при некотором $K > 0$. Устойчивость в форме неравенства $\|u^n\|^2 \leq \text{const} (\|u^0\|^2 + \|u^1\|^2)$ следует поэтому из рассуждений, аналогичных приведенным выше. Заметим, что в этом случае устойчивость определяется условием, вытекающим из требования положительной определенности S_n .

Если эта разностная схема применяется для конечного интервала $0 \leq x \leq 1$, то необходимо поставить дополнительное граничное условие при $x = 1$, чтобы иметь возможность решить систему разностных уравнений (6.30). Чтобы понять, как следует выбрать это граничное условие, сделаем пройденный выше путь в обратном направлении, считая, что $J \Delta x = 1$ и (6.30) выполняется для $j = 1, 2, \dots, J-1$; по этим же значениям j будет проводиться суммирование для норм и скалярных произведений. Теперь неравенство (6.33) больше не выполняется из-за наличия в левой части члена, содержащего u_j^n , и, чтобы сохранить (6.34), положим

$$S_n = \|u^n\|^2 + \|u^{n+1}\|^2 + 2(u^{n+1}, a \Delta_0 u^n) - a_{J-1} u_{J-1}^{n+1} u_J^n \Delta x.$$

Далее, мы видим, что вклад членов, содержащих u_J в выражении

$$2(u^{n-1}, a \Delta_0 u^n) + 2(u^n, a \Delta_0 u^{n-1}),$$

равен $a_{J-1} (u_{J-1}^{n-1} u_J^n + u_{J-1}^n u_J^{n-1}) \Delta x$. Поэтому вместо (6.32) будем иметь

$$S_n - S_{n-1} \leq L \Delta x (\|u^{n-1}\|^2 + \|u^n\|^2) - a_{J-1} u_J^n (u_{J-1}^{n-1} + u_{J-1}^{n+1}) \Delta x. \quad (6.35)$$

Ясно, что если положить

$$u_J^n = \frac{1}{2} (u_{J-1}^{n-1} + u_{J-1}^{n+1}), \quad (6.36)$$

то это не только позволит решить систему разностных уравнений, но также обеспечит и устойчивость: для последней точки сетки получается уравнение

$$u_{J-1}^{n+1} = \left(1 + \frac{1}{2} a_{J-1}\right)^{-1} \left[\left(1 - \frac{1}{2} a_{J-1}\right) u_{J-1}^n + a_{J-1} u_{J-2}^n \right], \quad (6.37)$$

которое обладает желательным свойством — имеет положительные коэффициенты, сумма которых равна единице.

Стоит отметить, что более непосредственный способ получения u_j^n экстраполяцией по j мог бы привести к неустойчивости. В этом случае вместо уравнения (6.37) мы имели бы

$$u_{J-1}^{n+1} = u_{J-1}^n - 2a_{J-1}(u_{J-1}^n - u_{J-2}^n). \quad (6.38)$$

Следовательно, если бы $a(x) \equiv 1$ и начальные данные имели вид

$$u_j^n = (-1)^{j+n}, \quad n = 0, 1, \quad j = 0, 1, \dots, J-1,$$

то вблизи правой границы решение можно было бы записать так:

$$u_j^n = (-1)^{j+n} + (-1)^j (j+n),$$

где $f(k+J) = 0$ при $k \leq 0$; подстановка этого выражения в (6.38) дает

$$f(k+J) = (-1)^{k-1} 2k(k+1) \quad \text{при } k > 0.$$

Энергетический метод может быть исключительно полезен для устранения неприятностей такого рода.

Прежде чем закончить рассмотрение простой задачи этого параграфа, следует заметить, что при замене $a(x)$ в уравнении (6.22) на $u(x, t)$ получается очень распространенный нелинейный оператор. Для аппроксимации $u \partial u / \partial x$ можно использовать как $\frac{1}{2}(\Delta x)^{-1} \Delta_0 u^2$, так и $(\Delta x)^{-1} u \Delta_0 u$, что с точки зрения точности одинаково хорошо. Однако с помощью энергетического метода иногда удастся найти такую линейную комбинацию этих разностных операторов, которая является более предпочтительной. Например, рассмотрим схему Кранка — Николсона

$$u_j^{n+1} + \frac{1}{2} Q(u_j^{n+1}) = u_j^n - \frac{1}{2} Q(u_j^n), \quad (6.39)$$

где $Q(u)$ — разностная аппроксимация для $(\Delta x) u \partial u / \partial x$. После возведения обеих частей уравнения (6.39) в квадрат и последующего суммирования получим

$$\begin{aligned} \|u^{n+1}\|^2 + \frac{1}{4} \|Q(u^{n+1})\|^2 - \left[\|u^n\|^2 + \frac{1}{4} \|Q(u^n)\|^2 \right] = \\ = -(u^{n+1}, Q(u^{n+1})) - (u^n, Q(u^n)). \end{aligned}$$

Если $Q(u) = 1/2 \Delta_0 u^2$ и граничными членами можно пренебречь, то

$$\left(u, \frac{1}{2} \Delta_0 u^2\right) = -\frac{1}{2} (u^2, \Delta_0 u),$$

тогда как при $Q(u) = u \Delta_0 u$

$$(u, u \Delta_0 u) = (u^2, \Delta_0 u).$$

Следовательно, полагая

$$Q(u) = \frac{1}{3} (\Delta_0 u^2 + u \Delta_0 u) = \frac{1}{3} (u + T_+ u + T_- u) \Delta_0 u, \quad (6.40)$$

получаем, что $(u, Q(u)) = 0$, и величина $\|u^n\|^2 + 1/4 \|Q(u^n)\|^2$ остается постоянной, но это и означает устойчивость.

§ 6.4. Одновременное распространение звука и тепла

Чтобы у читателя не создалось впечатления о чрезмерной простоте применения энергетического метода, рассмотрим сейчас более сложную систему уравнений. При этом временно будем пренебрегать эффектами, связанными с наличием конечных границ, и возьмем в качестве пространственного интервала всю действительную ось $-\infty < x < +\infty$. Постараемся также придать нашему изложению характер действительного проводимого на практике исследования, для чего используем вначале самые очевидные соображения и только позднее будем уточнять их по мере надобности.

Распространение звука (малых возмущений) в газе с теплопроводностью σ , изотермической скоростью звука c и адиабатической постоянной $\gamma > 1$ описывается системой уравнений

$$\begin{aligned} \frac{\partial u}{\partial t} &= c \frac{\partial}{\partial x} [\omega - (\gamma - 1) e], \\ \frac{\partial \omega}{\partial t} &= c \frac{\partial u}{\partial x}, \\ \frac{\partial e}{\partial t} &= \sigma \frac{\partial^2 e}{\partial x^2} - c \frac{\partial u}{\partial x}. \end{aligned} \quad (6.41)$$

Рассмотрим разностную аппроксимацию этой системы

$$\begin{aligned} u^{n+1} - u^n &= \nu [\Delta_+ \omega^n - (\gamma - 1) \Delta_+ e_n], \\ \omega^{n+1} - \omega^n &= \nu \Delta_- u^{n+1}, \\ e^{n+1} - e^n &= \mu \Delta_- \Delta_+ e^{n+1} - \nu \Delta_- u^{n+1}, \end{aligned} \quad (6.42)$$

где $\nu = c \Delta t / \Delta x$, $\mu = \sigma \Delta t / (\Delta x)^2$. Достаточно подробное изучение устойчивости этой схемы будет проведено в гл. 10, в основном с помощью метода рядов Фурье. Там будет показано, что при $\sigma \neq 0$ условием устойчивости является неравенство $\nu < 1$, то-

гда как при $\sigma = 0$ оно, как легко видеть, заменяется более сильным ограничением $\nu < \gamma^{-1/2}$. Это неожиданное изменение указывает на практическую ограниченность использовавшегося нами до сих пор понятия устойчивости, согласно которому влияние условия устойчивости на поведение решения рассматривается только в пределе при $\Delta t \rightarrow 0$. Мы еще вернемся к этому вопросу ниже при обсуждении энергетического метода, которое послужит дополнением к тому, что сказано по этому поводу в гл. 10.

В качестве первого шага при применении энергетического метода образуем скалярные произведения каждого разностного уравнения с суммой двух членов, стоящих в соответствующей левой части; в результате получим

$$\begin{aligned} \|u^{n+1}\|^2 - \|u^n\|^2 &= \nu(u^n + u^{n+1}, \Delta_+ w^n) - \nu(\gamma - 1)(u^n + u^{n+1}, \Delta_+ e^n), \\ \|w^{n+1}\|^2 - \|w^n\|^2 &= \nu(\Delta_- u^{n+1}, w^n + w^{n+1}) = -\nu(u^{n+1}, \Delta_+(w^n + w^{n+1})), \\ \|e^{n+1}\|^2 - \|e^n\|^2 &= \mu(\Delta_- \Delta_+ e^{n+1}, e^n + e^{n+1}) - \nu(\Delta_- u^{n+1}, e^n + e^{n+1}) = \\ &= -\mu(\Delta_+ e^{n+1}, \Delta_+(e^n + e^{n+1})) + \nu(u^{n+1}, \Delta_+(e^n + e^{n+1})). \end{aligned}$$

Умножая последнее из этих соотношений на $\gamma - 1$ и складывая затем их вместе, получим после приведения подобных членов соотношение, которое наводит на мысль определить S_n так:

$$\begin{aligned} S_n &= \|u^n\|^2 + \|w^n\|^2 + (\gamma - 1)\|e^n\|^2 + \\ &+ \frac{1}{2}(\gamma - 1)\mu\|\Delta_+ e^n\|^2 + \nu(u^n, \Delta_+(w^n - (\gamma - 1)e^n)). \end{aligned} \quad (6.43)$$

В этом случае

$$\begin{aligned} S_{n+1} - S_n &= -\frac{1}{2}(\gamma - 1)\mu[\|\Delta_+ e^n\|^2 + 2(\Delta_+ e^n, \Delta_+ e^{n+1}) + \\ &+ \|\Delta_+ e^{n+1}\|^2] \leq 0. \end{aligned} \quad (6.44)$$

Действительные трудности начинаются при попытке доказать, что S_n в самом деле определяет норму, когда $\nu < 1$, т. е. когда Δx задается соотношением $\nu = \text{const}$, так что $\mu = \text{const}(\Delta x)^{-1}$. Очевидно, что

$$|(u^n, \Delta_+ w^n)| \leq 2\|u^n\|\|w^n\| \leq \|u^n\|^2 + \|w^n\|^2 \quad (6.45)$$

и при любом $\alpha > 0$

$$|(u^n, \Delta_+ e^n)| \leq \|u^n\|\|\Delta_+ e^n\| \leq \frac{1}{2}(\alpha \Delta x \|u^n\|^2 + (1/\alpha \Delta x) \|\Delta_+ e^n\|^2). \quad (6.46)$$

Следовательно,

$$\begin{aligned} S_n &\geq (1 - \nu)\|w^n\|^2 + (\gamma - 1)\|e^n\|^2 + \\ &+ \left[1 - \nu\left(1 + \frac{1}{2}\alpha(\gamma - 1)\Delta x\right)\right]\|u^n\|^2 + \frac{1}{2}(\gamma - 1)(\mu - (\nu/\alpha \Delta x))\|\Delta_+ e^n\|^2. \end{aligned} \quad (6.47)$$

Последний член наводит на мысль положить $\alpha = \nu/\mu\Delta x = c/\sigma$, в этом случае он обращается в нуль. Однако, для того чтобы иметь положительный коэффициент при $\|u^n\|^2$, необходимо взять

$$\Delta x < [2\sigma/c\nu(\gamma - 1)](1 - \nu). \quad (6.48)$$

Таким образом, мы можем доказать устойчивость для произвольно малого отличного от нуля σ и для произвольно близких к 1 значений ν , но только при условии соответствующей малости интервала Δx .

Поскольку $S_n \leq S_{n-1} \leq \dots \leq S_1 \leq S_0$, то для завершения доказательства стандартным способом нужно было бы показать, что величина S_0 ограничена произведением некоторой постоянной на сумму $(\|u^0\|^2 + \|w^0\|^2 + \|e^0\|^2)$. Но сделать это нам мешает член $(\Delta x)^{-1}\|\Delta_+ e^0\|^2$. Поэтому вместо S_0 мы рассмотрим S_1 . Последнее из разностных уравнений (6.42) при $n = 1$ дает

$$\|e^1 - \mu\Delta_-\Delta_+ e^1\|^2 = \|e^0 - \nu\Delta_- u^1\|^2,$$

т. е.

$$\begin{aligned} \|e^1\|^2 + 2\mu\|\Delta_+ e^1\|^2 + \mu^2\|\Delta_-\Delta_+ e^1\|^2 = \\ = \|e^0 - \nu\Delta_- u^1\|^2 \leq 2\|e^0\|^2 + 8\nu^2\|u^1\|^2. \end{aligned}$$

Так как $\mu = (\sigma\nu/c)(\Delta x)^{-1}$, то это означает ограниченность члена $(\Delta x)^{-1}\|\Delta_+ e^1\|^2$; поэтому величина S_1 ограничена некоторой комбинацией квадратов L_2 -норм при $n = 0$ и $n = 1$. Теперь устойчивость получается обычным путем из оценки L_2 -нормы с помощью S_n .

С практической точки зрения эти рассуждения не имеют особой ценности, когда σ мало или когда требуется взять ν близким к 1, так как тогда условие (6.48) налагает слишком сильное ограничение на Δx . Что требуется на самом деле, так это точно оценить при любом заданном Δx то значение Δt , для которого скорость роста некоторой подходящей нормы решения остается ограниченной. В нашем случае последовательность S_n никогда не возрастает, и здесь необходимы более тонкие рассуждения для оценки ее нижней границы. Поэтому мы заменим (6.45) неравенством

$$|(u^n, \Delta_+ w^n)| \leq \beta\|u^n\|^2 + \beta^{-1}\|w^n\|^2,$$

где $\beta > 0$ будет выбрано позднее, и линейно скомбинируем (6.46) с аналогичной оценкой

$$|(u^n, \Delta_+ e^n)| \leq \delta\|u^n\|^2 + \delta^{-1}\|e^n\|^2$$

в отношении $(1 - \xi) : \xi$ где ξ и δ также будут выбраны позже. Теперь попытаемся получить оценку вида

$$S_n \geq a_1\|u^n\|^2 + a_2\|w^n\|^2 + a_3\|e^n\|^2 + a_4\|\Delta_+ e^n\|^2, \quad (6.49)$$

где a_1 , a_2 и a_3 строго положительны, а a_4 неотрицательно. Из предыдущего ясно, что

$$\begin{aligned} a_1 &= 1 - \nu [\beta + \xi (\gamma - 1) \delta + \frac{1}{2} (1 - \xi) (\gamma - 1) \alpha \Delta x], \\ a_2 &= 1 - \nu \beta^{-1}, \\ a_3 &= (\gamma - 1) (1 - \nu \xi \delta^{-1}), \\ a_4 &= \frac{1}{2} (\gamma - 1) [\mu - (1 - \xi) \nu \alpha^{-1} (\Delta x)^{-1}]. \end{aligned}$$

Для положительности последних трех выражений необходимо и достаточно, чтобы выполнялись неравенства

$$\beta > \nu, \quad \delta > \nu \xi \quad \text{и} \quad \alpha > c(1 - \xi)/\sigma.$$

При сопоставлении этих неравенств с выражением для a_1 становится ясным, что нужная нам оценка (6.49) имеет место в том случае, когда выполнено неравенство

$$\nu^2 [1 + (\gamma - 1) (\xi^2 + \frac{1}{2} (1 - \xi)^2 / \mu)] < 1,$$

левая часть которого имеет минимум при $\xi = (1 + 2\mu)^{-1}$. Следовательно, практическим критерием устойчивости является условие

$$\nu < \sqrt{(1 + 2\mu)(\gamma + 2\mu)}, \quad (6.50)$$

которое подтверждается численными экспериментами, описанными в гл. 10.

§ 6.5. Смешанные задачи для симметричных гиперболических систем

Чтобы проиллюстрировать применение энергетического метода в более общей ситуации, рассмотрим смешанную задачу для симметричной гиперболической системы, решение которой удовлетворяет некоторому энергетическому неравенству. Предположим, что комплекснозначный p -мерный вектор $u = u(x, t)$ есть решение системы

$$\frac{\partial u}{\partial t} = P(x, \partial/\partial x) u, \quad 0 \leq x \leq 1, \quad 0 \leq t \leq 1, \quad (6.51)$$

где матрица P представлена в виде

$$P(x, \partial/\partial x) u \equiv \frac{1}{2} \left[\frac{\partial}{\partial x} (A(x) u) + A(x) \frac{\partial}{\partial x} u \right] + B(x) u. \quad (6.52)$$

Здесь $A(x)$ и $B(x)$ — квадратные матрицы порядка p , гладко зависящие от x , но не зависящие от t , причем матрица $A(x)$ является эрмитовой. В качестве начального условия возьмем

$$u(x, 0) = f(x), \quad 0 \leq x \leq 1, \quad (6.53)$$

а граничные условия пусть заданы в виде

$$L(x)u(x, t) = 0, \quad x = 0, 1, \quad 0 \leq t \leq 1, \quad (6.54)$$

содержащем $\mu(x)$ линейно независимых однородных уравнений относительно компонент вектора u .

Предположим далее, что существует такая не зависящая от u постоянная K , для которой выполняется следующее *энергетическое неравенство*:

$$\frac{\partial}{\partial t} \|u(x, t)\|^2 \leq K \|u(x, t)\|^2, \quad 0 \leq t \leq 1. \quad (6.55)$$

Такое соотношение, включающее как частный случай закон сохранения энергии (равенство при $K = 0$), часто имеет место в физических задачах. Если решение u дифференцируемо, то из (6.55) следует, что

$$\frac{\partial}{\partial t} \|u\|^2 = \left(u, \frac{\partial u}{\partial t}\right) + \left(\frac{\partial u}{\partial t}, u\right) = 2 \operatorname{Re}(u, Pu) \leq K \|u\|^2. \quad (6.56)$$

Но так как

$$2 \operatorname{Re}(u, Pu) = u^* Au|_0 + 2 \operatorname{Re}(u, Bu),$$

то (6.56) сводится к условию в граничных точках

$$u^* Au|_0 \leq \operatorname{const} \|u\|^2. \quad (6.57)$$

Крайс [1960, 1963] назвал оператор, определяемый равенством

$$w = P(x, \partial/\partial x)v$$

для всех бесконечно дифференцируемых функций $v = v(x)$, удовлетворяющих граничным условиям (6.54), *полуограниченным*, если выполнено неравенство

$$2 \operatorname{Re}(v, Pv) \leq K \|v\|^2. \quad (6.58)$$

Предположим, что имеется несколько таких операторов, соответствующих одной и той же матрице P и различным наборам граничных условий. Тогда те из них, для которых полное число граничных условий $\mu(0) + \mu(1)$ минимально, называются *минимально полуограниченными операторами*. Крайс показал, что в случае, когда матрица P и граничные условия (6.54) определяют такой оператор, задача (6.51) — (6.54) поставлена корректно. Кроме того, граничные условия при этом должны удовлетворять следующим требованиям:

- 1) число строк $\mu(0)$ матрицы $L(0)$ должно быть равно числу отрицательных собственных значений матрицы $A(0)$;
- 2) число строк $\mu(1)$ матрицы $L(1)$ должно быть равно числу положительных собственных значений матрицы $A(1)$;

3) при $x = 0$ и $x = 1$ должно быть
 $(-1)^{x+1} v^* A v \leq 0$ для каждого вектора v , такого, что $L(x)v = 0$.
 (6.59)

Если матрица $A(x)$ имеет диагональную форму, к которой она может быть приведена с помощью унитарного преобразования $\hat{u} = Uu$, то эти требования можно сформулировать более явным образом. Предположим, что $A(x)$ и $u(x)$ при $x = 0$ и $x = 1$ представлены в виде

$$A(x) = \begin{bmatrix} \Lambda_1(x) & \Phi & \Phi \\ \Phi & -\Lambda_2(x) & \Phi \\ \Phi & \Phi & \Phi \end{bmatrix}, \quad u(x) = \begin{bmatrix} u_1(x) \\ u_2(x) \\ u_3(x) \end{bmatrix}, \quad (6.60)$$

где матрицы $\Lambda_1(x)$ и $\Lambda_2(x)$ диагональны, причем диагональные элементы положительны, Φ — нулевые матрицы и размерности всех матриц зависят от x . Тогда допустимые граничные условия имеют вид

$$u_2(0) = S(0)u_1(0), \quad u_1(1) = S(1)u_2(1), \quad (6.61)$$

где матрицы $S(0)$ и $S(1)$ удовлетворяют неравенствам

$$-\Lambda_1(0) + S^*(0)\Lambda_2(0)S(0) \leq 0, \quad (6.62)$$

$$-\Lambda_2(1) + S^*(1)\Lambda_1(1)S(1) \leq 0. \quad (6.63)$$

Теперь рассмотрим неявную разностную аппроксимацию уравнения (6.51)

$$(I - \alpha \Delta t P_\Delta) u_i^{n+1} = (I + (1 - \alpha) \Delta t P_\Delta) u_i^n, \quad i = 1, 2, \dots, J-1, \quad (6.64)$$

где $1/2 \leq \alpha \leq 1$, а через P_Δ обозначен разностный оператор, аппроксимирующий $P(x, \partial/\partial x)$. Зададим граничные условия в виде

$$\left. \begin{aligned} M(0) u^n &\equiv \sum_{j=0}^{J-1} m_j(0, \Delta x) u_j^n = 0, \\ M(1) u^n &\equiv \sum_{j=0}^{J-1} m_{J-j}(1, \Delta x) u_{J-j}^n = 0, \end{aligned} \right\} \quad (6.65)$$

где матричные коэффициенты $m_j(x, \Delta x)$ предполагаются подходящим образом нормированными. Напомним, что согласно общей теории, изложенной в гл. 3, граничные условия для задачи в дифференциальной форме учитываются при выборе того банахова пространства, которое используется в дальнейшем. Таким образом, для обеспечения согласованности приведенной выше аппроксимации мы потребуем, чтобы для любого решения u

дифференциального уравнения выполнялась импликация

$$L(x)u = 0 \text{ влечет } \|M(x)u\| \rightarrow 0 \text{ при } \Delta x \rightarrow 0, \text{ для } x = 0, 1. \quad (6.66)$$

Поэтому мы предположим, что число строк матрицы $M(x)$ меньше или равно числу строк матрицы $L(x)$. Упрощение возникает в том случае, когда отображение сеточной функции u^n в функцию u банахова пространства устроено так, что если сеточная функция удовлетворяет условию $M(x)u^n = 0$, то ее образ удовлетворяет условию $L(x)u = 0$. Например, когда точка $x = 0$ лежит посередине между точками сетки $j = 0, 1$, то граничное условие $u(0) = 0$ обычно аппроксимируют равенством $u_0^n + u_1^n = 0$; тогда отображение u^n в u с помощью линейной интерполяции обладает требуемым свойством.

Предположим дополнительно, что для любой последовательности значений сеточной функции u_j^n во внутренних точках значения u_0^n и u_J^n можно определить так, чтобы удовлетворялись граничные условия (6.65). Таким образом, последние не могут включать в себя никакого уравнения, содержащего лишь значения функции u_j^n во внутренних точках сетки. Мы можем также дополнить эти уравнения рядом других аналогичных им уравнений таким образом, чтобы u_0^n и u_J^n можно было однозначно вычислить по u_j^n , $j = 1, 2, \dots, J-1$. Аналогичные дополнительные уравнения могут быть использованы для всех других внешних точек сетки, которые входят в $P_\Delta u_j$, где j соответствует внутренней точке сетки.

Обозначим через $\mathcal{N}$ класс сеточных функций w , которые удовлетворяют граничным условиям (6.65) и этим дополнительным соотношениям. Тогда для устойчивости разностных уравнений можно получить условие, весьма близкое к тому, которое обеспечивает корректность постановки задачи в дифференциальной форме. А именно, *если существует такая постоянная K , что для всех достаточно малых Δx справедливо неравенство*

$$\operatorname{Re}(w, P_\Delta w)_\Delta \leq K \|w\|_\Delta^2, \text{ когда } w \in \mathcal{N}, \quad (6.67)$$

то при таких Δx и при $4K\Delta t < 1$ функции u^n , $n = 1, 2, \dots$, получаемые с помощью уравнения (6.64), однозначно определяются в классе $\mathcal{N}$ начальными условиями и разностная схема устойчива. Следует заметить, что норма $\|\cdot\|_\Delta$ строится только по внутренним точкам сетки.

Доказательство этого утверждения сводится к простому сопоставлению сделанных ранее предположений. Рассмотрим систему уравнений

$$(I - \alpha \Delta t P_\Delta) w_j = f_j, \quad j = 1, 2, \dots, J-1, \quad (6.68)$$

с произвольными правыми частями f_j . Если функция $w \in \mathcal{N}$, то согласно сказанному выше ее значения во внешних точках сетки можно исключить, после чего система (6.68) сведется к $p(J-1)$ линейным уравнениям относительно $p(J-1)$ неизвестных компонент функции w во внутренних точках. Далее, если $w \in \mathcal{N}$, $1/2 \leq \alpha \leq 1$ и $4K\Delta t < 1$, то

$$\begin{aligned} \|f\|_{\Delta}^2 &= \|(I - \alpha \Delta t P_{\Delta}) w\|_{\Delta}^2 \geq \|w\|_{\Delta}^2 - 2\alpha \Delta t \operatorname{Re}(w, P_{\Delta} w)_{\Delta} \geq \\ &\geq (1 - 2\alpha K \Delta t) \|w\|_{\Delta}^2 \geq 1/2 \|w\|_{\Delta}^2. \end{aligned} \quad (6.69)$$

Отсюда следует, что система (6.68) определяет единственную функцию $w \in \mathcal{N}$. Следовательно, для любой начальной сеточной функции $u^0 \in \mathcal{N}$ разностная схема (6.64) дает единственную последовательность $u^n \in \mathcal{N}$, $n = 1, 2, \dots$.

Чтобы доказать устойчивость, умножим скалярно обе части уравнения (6.64) на $w = \alpha u^{n+1} + (1 - \alpha)u^n$, что приводит к соотношению

$$\operatorname{Re}(w, u^{n+1} - u^n)_{\Delta} = \Delta t \operatorname{Re}(w, P_{\Delta} w)_{\Delta} \leq K \Delta t \|w\|_{\Delta}^2,$$

для левой части которого имеем

$$\begin{aligned} \alpha \|u^{n+1}\|_{\Delta}^2 - (2\alpha - 1) \operatorname{Re}(u^{n+1}, u^n)_{\Delta} - (1 - \alpha) \|u^n\|_{\Delta}^2 &\geq \\ &\geq [\alpha - 1/2(2\alpha - 1)] \|u^{n+1}\|_{\Delta}^2 - [1 - \alpha + 1/2(2\alpha - 1)] \|u^n\|_{\Delta}^2 = \\ &= 1/2 \|u^{n+1}\|_{\Delta}^2 - 1/2 \|u^n\|_{\Delta}^2. \end{aligned}$$

Но $\|w\|_{\Delta}^2 \leq \|u^{n+1}\|_{\Delta}^2 + \|u^n\|_{\Delta}^2$, так что

$$\|u^{n+1}\|_{\Delta}^2 - \|u^n\|_{\Delta}^2 \leq 2K \Delta t (\|u^{n+1}\|_{\Delta}^2 + \|u^n\|_{\Delta}^2),$$

откуда

$$\|u^n\|_{\Delta}^2 \leq (1 + 2K \Delta t)^n \cdot (1 - 2K \Delta t)^{-n} \|u^0\|_{\Delta}^2; \quad (6.70)$$

значит, разностная схема устойчива.

В качестве простого примера использования этих результатов применим схему Кранка — Николсона к изучавшейся в § 6.3 задаче

$$\frac{\partial u}{\partial t} + a(x) \frac{\partial u}{\partial x} = 0, \quad 0 \leq x \leq 1, \quad 0 \leq t \leq 1, \quad (6.71)$$

$$u(0, t) = 0, \quad u(x, 0) = F(x),$$

где функции u и a действительны и $a(x) > 0$. Сетка выбирается так, что точки $x = 0, 1$ отвечают значениям $j = 0, J$ соответственно, и схема Кранка — Николсона получается, если положить $\alpha = 1/2$ и

$$P_{\Delta} = -(\Delta x)^{-1} a(x) \Delta_0.$$

При $x = 1$ никакого граничного условия нет, а при $x = 0$

$$M(0) u^n \equiv u_0^n = 0.$$

Класс $\mathcal{N}$ характеризуется этим условием и теми дополнительными соотношениями, которые необходимы для получения неравенства (6.67). В силу (6.15) для любой сеточной функции w имеем

$$|(w, a\Delta_0 w) - (w, \Delta_0(aw))| \leq L \Delta x \|w\| \|w\|_{0,J},$$

где L — постоянная Липшица для функции $a(x)$. Далее, из тождества (6.19) следует что

$$(w, \Delta_0(aw)) = -(\Delta_0 w, aw) + \frac{1}{2} \Delta x [(a_J + a_{J-1}) w_J w_{J-1} - (a_0 + a_1) w_0 w_1].$$

Поэтому определим $\mathcal{N}$ равенством

$$\mathcal{N} = \{w \mid w_0 = 0, \quad w_J = w_{J-1} \text{ и } w_j = 0 \text{ при } j < 0 \text{ или } j > J\}, \quad (6.72)$$

так что

$$\operatorname{Re}(w, P_\Delta w) \leq L \|w\|^2 - \frac{1}{4} (a_J + a_{J-1}) (w_{J-1})^2 \leq L \|w\|^2 \text{ для } w \in \mathcal{N}.$$

Поэтому однозначная разрешимость и устойчивость рассматриваемой разностной схемы при $4L\Delta t < 1$ следуют из приведенной выше теоремы.

§ 6.6. Спектральный анализ и критерий устойчивости Годунова — Рябенского

В этом параграфе мы вернемся к исследованию локальных собственных колебаний как средству определения устойчивости. В случае переменных коэффициентов и граничных условий, не сводящихся к периодичности решения, мы не можем, конечно, непосредственно применить преобразование Фурье и свести дело к алгебраической задаче, изучавшейся в гл. 4. Даже когда пространство действительно, использование конечноразностной аппроксимации для вычислений на каждом шаге по времени во всех тех случаях, которые мы до сих пор рассматривали, сводится к решению системы линейных алгебраических уравнений; эта система уравнений должна обладать всеми свойствами, необходимыми для выяснения вопроса об устойчивости соответствующей разностной схемы. Основная трудность состоит в том, что при $\Delta t \rightarrow 0$ порядок этой системы возрастает. Вследствие этого, например, может оказаться, что собственные значения соответствующих матриц могут только ввести в заблуждение при изучении устойчивости.

К примеру, возьмем такую аппроксимацию нашей модельной задачи (6.71):

$$u_j^{n+1} = (1 - r_j) u_j^n + r_j u_{j-1}^n, \quad r_j = (\Delta t / \Delta x) a_j. \quad (6.73)$$

Тогда, если рассматривать $\{u_j^n, j=0, 1, \dots, J\}$ как вектор, то переход к следующему временному слою будет описываться матрицей

$$C = \begin{bmatrix} 0 & & & & & \\ r_1 & 1-r_1 & & & & 0 \\ & r_2 & 1-r_2 & & & \\ & & \ddots & \ddots & & \\ 0 & & & & \ddots & \\ & & & r_J & 1-r_J & \end{bmatrix}. \quad (6.74)$$

Ее собственные значения равны 0 и $1-r_j, j=1, 2, \dots, J$, и лежат внутри единичной окружности при $0 < r_j < 2$; таким образом, при каждом фиксированном Δt ошибки в конце концов затухают (некоторые авторы берут это свойство в качестве определения устойчивости). Однако собственно устойчивость определяется хорошо известным условием Куранта — Фридрихса — Леви $0 < r_j < 1$; вне этого интервала сходимости при $\Delta t \rightarrow 0$ нет. Имеется различие между тем, что происходит при $n \rightarrow \infty$ и фиксированном Δt , и требуемым поведением при $n \rightarrow \infty$ и фиксированном значении $n\Delta t$. Например, если все $r_j = 3/2$ и u_j^0 лишь меняет знак от точки к точке, то значение $|u_j^n|$ в каждой точке сетки сначала возрастает в соответствии с формулой $u_j^n = (-2)^n u_j^0$, до тех пор пока до этой точки не дойдет влияние граничного условия, которое является единственной причиной окончательного затухания; но при $\Delta t \rightarrow 0$ этот начальный рост оказывается бесконечным (по поводу дальнейших примеров такого рода см. Партер [1962]).

Значительно лучше характеризуется устойчивость L_2 -нормой матрицы C . Рассматривая векторы с знакопередающимися компонентами, легко показать, что если все $r_j = r$, то

$$\|C\| = r + |1-r| + O(1/J);$$

действительно, спектральный радиус бесконечной матрицы вида (6.74) равен $r + |1-r|$.

Соображения такого рода побудили Годунова и Рябенского [1963а] ввести понятие спектра семейства операторов.

Пусть $\mathcal{F} = \{C(\Delta t) | \Delta t > 0\}$ — однопараметрическое семейство линейных операторов в банаховом пространстве $\mathcal{B}$. Комплексное число λ называется точкой спектра семейства $\mathcal{F}$, если существуют такие последовательности $\Delta t^{(J)}$ и $u^{(J)} \in \mathcal{B}$, что $\|u^{(J)}\| = 1, \Delta t^{(J)} \rightarrow 0$ и

$$\|[C(\Delta t^{(J)} - \lambda I)]u^{(J)}\| \rightarrow 0 \text{ при } J \rightarrow \infty. \quad (6.75)$$

Если семейство состоит из одного оператора, то это определение совпадает с обычным. Пример семейства (6.74) показывает, что это определение является нетривиальным обобщением; действительно, как доказали Годунов и Рябенский, в том случае, когда все $r_j = r$, спектр семейства представляет собой замкнутый круг радиуса r с центром в точке $(1 - r, 0)$. Для того чтобы убедиться в этом, достаточно рассмотреть при $\lambda = 1 - r + s$ и $|s| \leq r$ векторы, компоненты которых получаются усечением и последующей нормировкой последовательности $(s/r)^{-j}$, $j = 0, 1, \dots$.

Ясно, что необходимое условие устойчивости семейства разностных операторов состоит в том, что его спектр должен лежать в замкнутом единичном круге. Это требование мы будем называть *критерием устойчивости Годунова — Рябенского*. Очевидна его аналогия с условием фон Неймана. При применении критерия Годунова — Рябенского в качестве $u^{(j)}$ в (6.75) подставляются все локальные собственные колебания с целью проверить, удовлетворяют ли они этому критерию. Так, при рассмотрении локальных колебаний Фурье получается условие на собственные значения λ матрицы перехода; в пределе при $\Delta t \rightarrow 0$ они будут, очевидно, принадлежать спектру семейства и, следовательно, должно выполняться условие $|\lambda| \leq 1 + O(1)$. Это условие значительно слабее условия фон Неймана $|\lambda| \leq 1 + O(\Delta t)$. Но с другой стороны, критерий Годунова — Рябенского является более сильным в том смысле, что при этом рассматривается более широкий класс собственных колебаний. Например, вблизи границы вполне могут существовать колебания, которые затухают вдали от нее и оказывают влияние на устойчивость разностной схемы.

Простым примером может служить волновое уравнение на полуоси $0 \leq x < \infty$, которое мы запишем в виде системы

$$\frac{\partial u}{\partial t} = c \frac{\partial v}{\partial x}, \quad \frac{\partial v}{\partial t} = c \frac{\partial u}{\partial x} \quad (6.76)$$

и аппроксимируем разностной схемой

$$\begin{aligned} u_j^{n+1/2} - u_j^{n-1/2} &= r \Delta_+ v_{j-1/2}^n, \\ v_{j+1/2}^{n+1} - v_{j+1/2}^n &= r \Delta_+ u_j^{n+1/2}, \quad j, n = 0, 1, 2, \dots, \end{aligned} \quad (6.77)$$

где $r = c\Delta t/\Delta x$; рассмотрение матрицы перехода для этой схемы приводит к условию устойчивости $r < 1$ (см. гл. 10). В качестве начальных условий задаются значения $u_j^{-1/2}$ и v_j^0 , и необходимо лишь одно граничное условие при $x = 0$, общий вид которого (с учетом того, что u и v могут меняться ролями) дается равенством

$$v_{-1/2}^{n+1} + \alpha u_0^{n+1/2} = 0, \quad (6.78)$$

где α — некоторая постоянная.

Для нахождения собственных колебаний ищем решение в виде

$$\tilde{w}_j^n = (\lambda)^n (\mu)^j \tilde{w}, \quad (6.79)$$

где через w_j^n обозначен вектор $(u_j^{n-1/2}, v_{j-1/2}^n)^T$, а $\tilde{w} = (\tilde{u}, \tilde{v})^T$ — постоянный вектор. Подставляя такое решение в (6.77), получаем

$$\begin{aligned} (\lambda - 1) \tilde{u} - r(\mu - 1) \tilde{v} &= 0, \\ (\lambda - 1) \mu \tilde{v} - r\lambda(\mu - 1) \tilde{u} &= 0; \end{aligned} \quad (6.80)$$

приравнивание нулю определителя этой системы из двух уравнений дает

$$\mu^2 - \left[2 + \frac{(\lambda - 1)^2}{r^2 \lambda} \right] \mu + 1 = 0. \quad (6.81)$$

Формула (6.79) дает надлежащее решение при $|\mu| \leq 1$, а неустойчивость возникает лишь при $|\lambda| > 1$.

Но если $r < 1$, что мы и будем предполагать, то решения с $|\mu| = 1$ имеют $|\lambda| = 1$ и потому устойчивы. Таким образом, в области $|\lambda| > 1$ решение уравнения (6.81), а именно

$$\mu = 1 + \frac{(\lambda - 1)^2}{2r^2 \lambda} \pm \sqrt{\frac{(\lambda - 1)^2}{r^2 \lambda} + \frac{(\lambda - 1)^4}{4r^4 \lambda^2}} \quad (6.82)$$

распадается на две ветви, для одной из которых $|\mu| > 1$, а для другой $|\mu| < 1$. Нас интересует последняя, поэтому мы возьмем в (6.82) знак минус и ту ветвь корня, которая положительна при скажем $\lambda > 1$.

Подстановка этого значения μ в уравнение (6.80) показывает, что можно взять, например,

$$\begin{aligned} \tilde{u} &= (\lambda - 1) - \sqrt{(\lambda - 1)^2 + 4\lambda r^2}, \\ \tilde{v} &= 2\lambda r. \end{aligned} \quad (6.83)$$

Тогда из граничного условия $\tilde{v} + \alpha \tilde{u} = 0$ получаем

$$\lambda = (\alpha + r\alpha^2)(\alpha + r). \quad (6.84)$$

Теперь нетрудно убедиться, что задача в дифференциальной форме поставлена корректно, если $\alpha \neq 1$ ¹⁾. Поведение функции

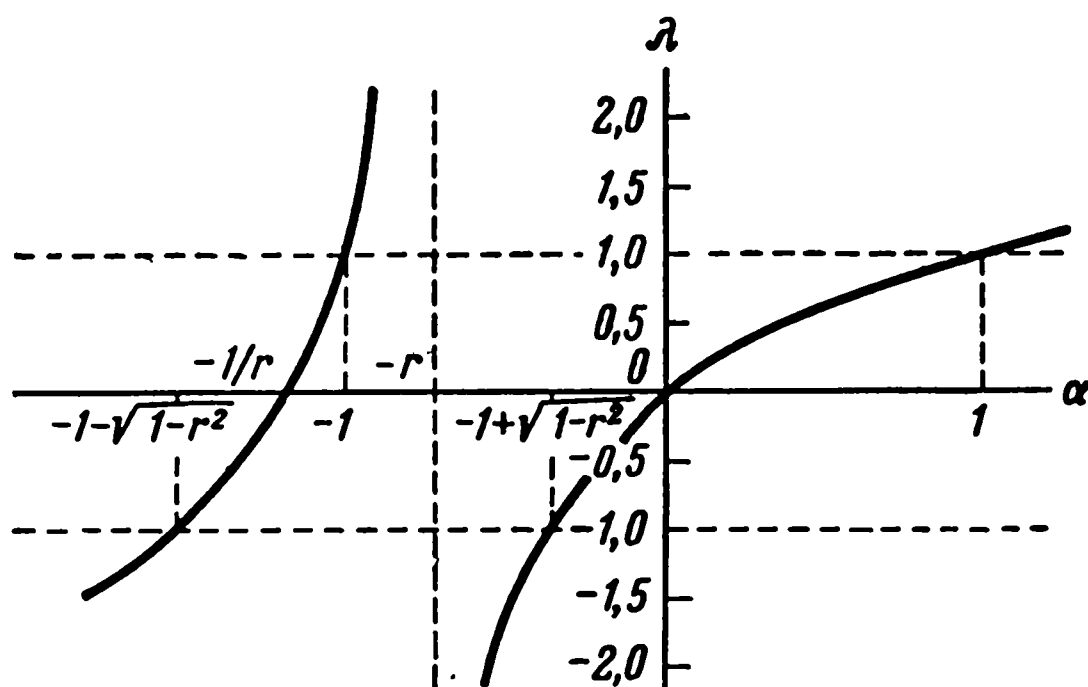
¹⁾ Стоит отметить, что приведенные в § 6.5 достаточные для корректности постановки дифференциальной задачи условия сводятся в силу соотношения (6.59) к неравенству $\tilde{u}\tilde{v} \geq 0$ при $x = 0$; это означает, что при $\alpha \leq 0$ задача поставлена корректно. Чтобы получить весь набор возможных граничных условий, нужно взять в качестве аналога энергии другую величину, или, что то же самое, перейти к новым переменным. Так, полагая $\hat{u} = \tilde{u} + \tilde{v}$ и $\hat{v} = \beta(\tilde{u} - \tilde{v})$ при $0 < \beta < \infty$, мы приведем допустимые граничные условия к виду $\hat{v} = \gamma \hat{u}$ при $x = 0$, где $|\gamma| \leq 1$. Это охватывает все случаи равенства $\tilde{v} + \alpha \tilde{u} = 0$ за исключением $\alpha = 1$.

(6.84) схематически изображено на рис. 6.1. Из рисунка видно, что (6.79) представляет собой неустойчивое колебание за исключением тех случаев, когда

$$1 > \alpha \geq -1 + \sqrt{1-r^2} \quad \text{или} \quad -1 \geq \alpha \geq -1 - \sqrt{1-r^2};$$

значения α вне этих интервалов приводят к неустойчивости. Отметим, что обычное граничное условие $\tilde{u} - \tilde{v} = 0$ является предельным случаем допустимых граничных условий.

В нашем примере $(C - \lambda I)u$ действительно обращается в нуль для рассмотренных выше колебаний, но в случае конечного



Р и с. 6.1. График функции (6.84).

интервала $0 \leq x \leq 1$ может оказаться, что u будет удовлетворять граничному условию в точке $x = 1$ лишь в пределе при $\Delta t \rightarrow 0$, т. е. при $J \rightarrow \infty$; например, это имеет место для той модельной задачи, которая приводится в начале данного параграфа. Далее, если уравнение (6.81) имеет кратные корни, то необходимо рассматривать колебания более сложного по сравнению с (6.79) вида, содержащие полиномы относительно j .

§ 6.7. Применение критерия Годунова — Рябенского к смешанным задачам

Рассмотрим на интервале $0 \leq x \leq 1$ общую двухслойную систему разностных уравнений вида

$$\sum_{k=0}^K B_k u_{j+k}^{n+1} = \sum_{k=0}^K A_k u_{j+k}^n, \quad j=0, 1, \dots, J-K, \quad n=0, 1, 2, \dots, \quad (6.85)$$

здесь u_j^n — это, как обычно, p -мерный вектор, а A_k и B_k — квадратные матрицы порядка p . На двух концах интервала заданы

граничные условия

$$\sum_{k=0}^{k_0} \beta_{0k} u_k^{n+1} = \sum_{k=0}^{k_0} \alpha_{0k} u_k^n, \quad (6.86)$$

$$\sum_{k=0}^{k_1} \beta_{1k} u_{j-k}^{n+1} = \sum_{k=0}^{k_1} \alpha_{1k} u_{j-k}^n, \quad (6.87)$$

где α_{ik} и β_{ik} — прямоугольные матрицы.

Чтобы продемонстрировать применение критерия Годунова — Рябенского к этой системе уравнений, удобно (хотя в данном конкретном случае и не обязательно) привести ее к такому специальному виду, когда $K = 1$ и $k_0 = k_1 = 0$. Это можно сделать следующим образом. Возьмем интервал I_0 на оси x , который начинается в самой левой точке сетки и идет вправо до l -й точки сетки, где $l \geq \max(k_0 + 1, k_1 + 1, K)$. Затем построим интервалы $I_1, I_2, \dots$, последовательно перемещая I_0 вправо на одну точку сетки до тех пор, пока не покроем все точки сетки. Теперь все компоненты векторов u_k^n , нижние индексы которых соответствуют интервалу I_m , объединим в новый вектор U_m^n , который тем самым будет иметь pl компонент. В результате система (6.86) преобразуется в систему однородных уравнений относительно U_0^n и U_0^{n+1} и аналогично преобразуется система (6.87). Компоненты векторов $U_m^n, U_{m+1}^n, U_m^{n+1}$ и U_{m+1}^{n+1} связаны системой pl скалярных уравнений, получающейся путем объединения p уравнений (6.85) при каждом из l возможных значений j , и системой $p(l-1)$ тождественных соотношений, характеризующих тот факт, что интервалы I_m и I_{m+1} имеют $l-1$ общих точек. Однако поскольку в общем случае число возможных значений m меньше числа значений j , для которых выполняется (6.85), то часть уравнений останется неиспользованной; эти уравнения относятся к последним интервалам и должны быть присоединены к граничным условиям. Как читатель заметит в следующей главе, способ сведения многослойной схемы к двухслойной очень напоминает то, что было сделано выше. При рассмотрении многослойного варианта данной задачи можно одновременно осуществить две редукции, используя вместо интервалов I_m прямоугольники в плоскости x, t .

Предположим теперь, что эта редукция выполнена, и рассмотрим задачу (6.85) — (6.87) с $K = 1$ и $k_0 = k_1 = 0$. Годунов и Рябенский [1963а] предполагали, что матричные коэффициенты A_k и B_k кусочно непрерывны по x и непрерывны по Δx в некотором интервале $0 \leq \Delta x \leq \tau$. Однако для простоты мы рассмотрим только случай постоянных матриц; зависимость от Δx можно учесть очень легко, а на вопросе о влиянии зависимости от x мы остановимся в конце параграфа.

Предположим, что при заданных u_j^n ($j=0, 1, \dots, J$) значения u_j^{n+1} однозначно определяются с помощью наших уравнений. Таким образом, сумма числа ненулевых строк матриц β_{00} и β_{10} равна p и существует такая матрица C_J порядка $p(J+1)$, что

$$\{u_j^{n+1} | j=0, 1, \dots, J\} = C_J \{u_j^n | j=0, 1, \dots, J\}.$$

К этому семейству разностных операторов $\{C_J\}$ мы и хотим применить критерий Годунова — Рябенского.

Точка λ комплексной плоскости, не принадлежащая спектру семейства, называется *регулярной*. Ясно, что для каждой такой точки при всех достаточно больших J существуют резольвенты $(C_J - \lambda I)^{-1}$, нормы которых ограничены константой, не зависящей от J . Ввиду приведенного выше уравнения это означает, что если мы положим

$$\begin{aligned} D_i &= \lambda B_i - A_i, \quad i=0, 1, \\ \delta_i &= \lambda \beta_{i0} - \alpha_{i0}, \quad i=0, 1, \end{aligned} \quad (6.88)$$

то для достаточно больших J система уравнений

$$\begin{aligned} D_0 u_j + D_1 u_{j+1} &= f_{j+1/2}, \quad j=0, 1, \dots, J-1, \\ \delta_0 u_0 &= \varphi, \quad \delta_1 u_J = \psi \end{aligned} \quad (6.89)$$

при любых правых частях имеет единственное решение, для которого справедлива оценка

$$\max_j |u_j| \leq M \max_j (|f_{j+1/2}|, |\varphi|, |\psi|), \quad (6.90)$$

где M не зависит от $f_{j+1/2}$, φ , ψ и J ; в этом случае система (6.89) называется *хорошо обусловленной*. Таким образом, мы можем сказать, что *разностная схема удовлетворяет критерию Годунова — Рябенского, который необходим для устойчивости, если для всех точек λ , лежащих вне единичного круга, уравнения (6.89) хорошо обусловлены*.

Для получения признаков того, что данные уравнения хорошо обусловлены, нам потребуется следующая лемма, принадлежащая Рябенскому [1964].

Пусть D_0 и D_1 — такие квадратные матрицы порядка p , что $\det(D_0 + \mu D_1) \neq 0$ при всех μ с $|\mu| = 1$, и пусть μ_0 есть k -кратный нуль для $\det(D_0 + \mu D_1)$, причем $|\mu_0| < 1$. Тогда существует такое k -мерное подпространство $\mathcal{R}(\mu_0)$ p -мерного векторного пространства, что если $V_0 \in \mathcal{R}(\mu_0)$, то уравнения

$$D_0 V_j + D_1 V_{j+1} = 0, \quad j=0, 1, 2, \dots, \quad (6.91)$$

имеют единственное решение $\{V_j\}$, для которого $|V_j| \rightarrow 0$ при $j \rightarrow \infty$; кроме того найдется такое $\alpha > 0$, что $|V_j| \leq \text{const } e^{-\alpha j}$ при $j=0, 1, 2, \dots$.

Доказательство. По предположению матрица $D_0 + D_1$ не вырождена, так что можно написать

$$D_0 + \mu D_1 = (D_0 + D_1) [I + (\mu - 1) E],$$

где $E = (D_0 + D_1)^{-1} D_1$. Если, далее, Q — жорданова каноническая форма матрицы E и $E = P Q P^{-1}$, то

$$D_0 + \mu D_1 = (D_0 + D_1) P [I + (\mu - 1) Q] P^{-1} \quad (6.92)$$

и μ_0 есть k -кратный нуль для $\det [I + (\mu - 1) Q]$. Матрица этого определителя имеет верхнюю треугольную форму, так что определитель равен $\prod_i [1 + (\mu - 1) v_i]$, где v_i — собственные значения матрицы E . В точности k из этих множителей должны

быть пропорциональны $\mu - \mu_0$; это значит, что $1 + (\mu - 1) v_i = v_i (\mu - \mu_0)$ и, следовательно, $v_i = (1 - \mu_0)^{-1}$ при k значениях i .

Поэтому мы можем считать, что строки и столбцы матрицы Q упорядочены таким образом, что первые k ее диагональных элементов равны $(1 - \mu_0)^{-1}$. Теперь возьмем первые k столбцов матрицы P в качестве базиса подпространства $\mathcal{R}(\mu_0)$ и запишем нужную нам последовательность векторов этого подпространства в виде

$$V_j = P a_j, \quad (6.93)$$

где a_j — векторы-столбцы, все компоненты которых кроме первых k равны нулю. Из тождества (6.92) относительно μ следует, что

$$D_0 P = (D_0 + D_1) (P - P Q) \quad \text{и} \quad D_1 P = (D_0 + D_1) P Q;$$

поэтому уравнения (6.91) принимают вид

$$(D_0 + D_1) P [I a_j - Q (a_j - a_{j+1})] = 0, \quad j = 0, 1, 2, \dots \quad (6.94)$$

При любом начальном векторе $V_0 \in \mathcal{R}(\mu_0)$, порождаемом начальным вектором a_0 , эти уравнения могут быть решены относительно последовательности a_j , так как в них входит только левый верхний минор порядка k матрицы Q , который отличен от нуля. Обозначая его через $\hat{Q}$ и используя аналогичные обозначения для I и a_j , получаем

$$\hat{a}_j = [\hat{Q}^{-1} (\hat{Q} - \hat{I})]' \hat{a}_0.$$

Матрица в квадратных скобках имеет верхнюю треугольную форму, и ее диагональные элементы равны μ_0 ; следовательно, элементы j -й степени этой матрицы ограничены модулем произведения μ_0^j на полином степени k от j , т. е. величиной $\text{const } e^{-\alpha j}$, где $|\mu_0| < e^{-\alpha} < 1$. Очевидно, таким же образом ограничены и компоненты векторов V_j , определяемых равенством (6.93).

Обозначим через $\mathcal{R}$ прямую сумму всех подпространств типа $\mathcal{R}(\mu_0)$, т. е.

$$\mathcal{R} = \Sigma_{\oplus} \{ \mathcal{R}(\mu_i) \mid \det(D_0 + \mu_i D_1) = 0, |\mu_i| < 1 \}. \quad (6.95)$$

Аналогично предыдущему существуют подпространства $\mathcal{S}(\eta_i)$, соответствующие таким корням η_i уравнения $\det(\eta D_0 + D_1) = 0$, для которых $|\eta_i| < 1$. Последовательность решений уравнения (6.91), начальный вектор которой принадлежит одному из этих подпространств, стремится к нулю при $j \rightarrow -\infty$; обозначим прямую сумму таких подпространств через $\mathcal{S}$, т. е.

$$\mathcal{S} = \Sigma_{\oplus} \{ \mathcal{S}(\eta_i) \mid \det(\eta_i D_0 + D_1) = 0, |\eta_i| < 1 \}. \quad (6.96)$$

Так как в силу равенства (6.92) уравнение $\det(D_0 + \mu D_1) = 0$ эквивалентно уравнению

$$\det[I + (\mu - 1)Q] = \prod_i [1 + (\mu - 1)v_i] = 0,$$

то для перехода от μ к η можно воспользоваться матрицей $I + (\mu - 1)Q$. Отсюда ясно, что если $\det D_1 = 0$, то уравнение $\det(\eta D_0 + D_1) = 0$ должно иметь внутри единичного круга помимо отличных от нуля корней η_i , соответствующих бóльшим единицы по модулю корням уравнения $\det(D_0 + \mu D_1) = 0$, также и нулевые корни, причем сумма числа таких корней и степени многочлена $\det(D_0 + \mu D_1)$ должна быть равна p . Поскольку по предположению $\det(D_0 + \mu D_1)$ не обращается в нуль при $|\mu| = 1$, то полное число корней μ_i и η_i внутри единичного круга также равно p . Это означает, что прямая сумма подпространств $\mathcal{R}$ и $\mathcal{S}$ совпадает со всем p -мерным пространством.

Мы уже убедились в том, что при $V_0 \in \mathcal{R}(\mu_0)$ решение уравнения (6.91) единственно в $\mathcal{R}(\mu_0)$. Ясно также, что единственность имеет место для всего подпространства $\mathcal{R}$, так как главный минор матрицы Q , соответствующий всем корням μ_i , также отличен от нуля. Однако если матрица D_1 вырождена, то к любому вектору V_{j+1} можно добавить такой вектор $W \in \mathcal{S}$, для которого $D_1 W = 0$. Но последовательность решений уравнения (6.91) с таким начальным вектором либо перестает быть непрерывной после конечного числа шагов, либо экспоненциально возрастает; поэтому найденное выше убывающее решение является единственным во всем пространстве и лемма Рябенского доказана.

Теперь мы можем сформулировать и доказать основной результат Годунова и Рябенского:

Для того чтобы система уравнений (6.89) была хорошо обусловлена, необходимо и достаточно выполнение следующих условий:

$$(1) \det(D_0 + \mu D_1) \neq 0 \text{ при } |\mu| = 1;$$

(2a) размерность подпространства $\mathcal{R}$ равна числу строк матрицы δ_0 , и для $v \in \mathcal{R}$ из равенства $\delta_0 v = 0$ следует, что $v = 0$;

(2b) размерность подпространства $\mathcal{S}$ равна числу строк матрицы δ_1 , и для $w \in \mathcal{S}$ из равенства $\delta_1 w = 0$ следует, что $w = 0$.

Здесь $\mathcal{R}$ и $\mathcal{S}$ — подпространства, определенные выше¹⁾.

Докажем сначала необходимость этих условий. Для этого мы покажем, что если какое-либо из них не выполняется, то можно найти последовательность таких векторов u_j с $\max_j |u_j| = 1$, что правые части уравнений (6.89) могут быть сделаны сколь угодно малыми при достаточно больших J , и таким образом получим противоречие с условием (6.90). Прежде всего, если $\det(D_0 + \mu_0 D_1) = 0$ при $|\mu_0| = 1$, то можно найти такой вектор v , что $(D_0 + \mu D_1)v = 0$ и $|v| = 1$, и положить $u_j = 4j(J-j)J^{-2}\mu_0^j v$. Тогда $\varphi = \psi = 0$ и $f_{j+1/2} = O(1/J)$. Далее, если размерность $\mathcal{R}$ больше числа строк матрицы δ_0 , то найдется такой вектор $u_0 \in \mathcal{R}$, что $|u_0| = 1$ и $\delta_0 u_0 = 0$. Тогда последовательность решений уравнений $D_0 u_j + D_1 u_{j+1} = 0$, $j = 0, 1, \dots, J-1$, дает $\varphi = 0$, $f_{j+1/2} = 0$, и $\psi = O(e^{-\alpha j})$ для некоторого $\alpha > 0$. Наконец, если размерность $\mathcal{R}$ меньше числа строк матрицы δ_0 , то размерность $\mathcal{S}$ должна быть больше числа строк матрицы δ_1 и потому не выполняется условие (2b).

Для доказательства достаточности этих условий нужно найти соответствующим образом ограниченное решение системы уравнений (6.89) в каждом из следующих случаев:

(a) $\varphi = 0$, $\psi = 0$ и $f_{j+1/2} = 0$ при $j \neq j_0$;

(b) $\varphi = 0$, все $f_{j+1/2} = 0$;

(c) $\psi = 0$, все $f_{j+1/2} = 0$.

В случае (a) заметим сначала, что существуют единственные векторы $v \in \mathcal{R}$ и $w \in \mathcal{S}$, такие, что $D_0 w + D_1 v = f_{j_0+1/2}$. Действительно, $f_{j_0+1/2} = f_0 + f_1$, где $f_0 \in \mathcal{R}$ и $f_1 \in \mathcal{S}$, причем это разложение единственно; v и w суть единственные решения (каждое в соответствующем подпространстве) уравнений $D_1 v = f_0$ и $D_0 w = f_1$. Теперь построим последовательность u_j , взяв в качестве начального вектора u_{j_0} сперва w , а затем v и решая уравнение $D_0 u_j + D_1 u_{j+1} = 0$ соответственно при $j < j_0$ и при $j > j_0$. В результате получим, что для $j \geq j_0 + 1$ $u_j \in \mathcal{R}$ и $|u_j|$ убывает при возрастании j , а для $j \leq j_0$ $u_j \in \mathcal{S}$ и $|u_j|$ убывает при убывании j .

Эта последовательность удовлетворяет основным уравнениям системы (6.89), но не удовлетворяет граничным условиям.

¹⁾ Напомним, что по предположению D_0 и D_1 не зависят от J , так что $\mathcal{R}$, $\mathcal{S}$ и используемая ниже постоянная α также не зависят от J .

Чтобы удовлетворить им, добавим сходящийся ряд поправок, т. е. представим полное решение в виде

$$u_j + v_j^{(1)} + w_j^{(1)} + v_j^{(2)} + w_j^{(2)} + \dots \quad (6.97)$$

Здесь вектор $v_0^{(1)}$ из $\mathcal{R}$ выбирается таким образом, чтобы $\delta_0(u_0 + v_0^{(1)}) = 0$; и этот выбор единствен в силу предположения (2a); аналогично $w_j^{(1)} \in \mathcal{S}$ выбирается так, чтобы $\delta_1(u_j + w_j^{(1)}) = 0$, а вектор $v_0^{(2)} \in \mathcal{R}$ так, чтобы $\delta_0(w_0^{(1)} + v_0^{(2)}) = 0$, и т. д. Далее, векторы $v_j^{(i)} \in \mathcal{R}$ выбираются так, чтобы $D_0 v_j^{(i)} + D_1 v_{j+1}^{(i)} = 0$, а векторы $w_j^{(i)} \in \mathcal{S}$ так, чтобы удовлетворялось то же соотношение $D_0 w_j^{(i)} + D_1 w_{j+1}^{(i)} = 0$. Так как $|v_j^{(i)}| \leq \leq \text{const } e^{-\alpha j} |v_0^{(i)}|$ и $|w_j^{(i)}| < \text{const } e^{-\alpha j} |w_0^{(i)}|$, то ясно, что при достаточно больших J ряд поправок сходится и тем самым граничные условия будут удовлетворены.

Решения системы уравнений (6.89) в оставшихся случаях (b) и (c) получаются аналогичным образом. Нам остается показать, что общее решение, являющееся линейной комбинацией этих фундаментальных решений, удовлетворяет условию ограниченности (6.90), несмотря на то, что число членов в этой комбинации возрастает с ростом J . Искомую постоянную M можно найти, поскольку сильно отличающиеся друг от друга значения j мало влияют друг на друга. Фундаментальное решение экспоненциально убывает при удалении j от j_0 , так что решение в точке сетки j_0 , получающееся при учете всех правых частей $f_{j+1/2}$, ограничено величиной

$$\text{const} \sum_{j=-\infty}^{\infty} \exp(-\alpha |j - j_0|),$$

а этот ряд сходится.

Мы определили хорошую обусловленность системы (6.89) и сформулировали определяющие это свойство условия, используя в качестве нормы максимум модуля, потому что она, по-видимому, является наиболее удобной и подходящей. Однако стоит отметить, что приведенные выше рассуждения остаются в силе и при использовании L_2 -нормы $\|u\| = \left[(1/J) \sum_j |u_j|^2 \right]^{1/2}$.

В этом случае (6.90) заменяется неравенством

$$\|u\| \leq \max \{ \|f\|, |\phi|, |\psi| \}, \quad (6.98)$$

где M снова не зависит от J . При доказательстве необходимости условий теоремы построение последовательности векторов в случае корня μ_0 с единичным модулем опять приводит к противоречию с (6.98); если же размерность $\mathcal{R}$ больше числа строк матрицы δ_0 , то, выбирая в качестве начального такой вектор $u_0 \in \mathcal{R}$, для которого $|u_0| = 1$ и $\delta_0 u_0 = 0$, получим, что

$\|u\| \geq J^{-1/2}$, $\|f\| = 0$, $|\varphi| = 0$ и $|\psi| = O(e^{-\alpha J})$, и потому неравенство (6.98) опять не выполняется. При доказательстве достаточности необходимые изменения также почти очевидны.

Этим завершается доказательство теоремы. При практическом применении критерия Годунова — Рябенского можно работать непосредственно с уравнениями в общем виде (6.85), рассматривая например корни уравнения $\det(D_0 + \mu D_1 + \dots + \mu^k D_k) = 0$, где $D_k = \lambda B_k - A_k$. Подпространства $\mathcal{R}$ и $\mathcal{P}$ обычно зависят не только от λ , но также от Δt или Δx . Однако если заранее известно, что матрицы A_k и B_k непрерывно зависят от Δt , то условия теоремы достаточно проверять только при $\Delta t = 0$. Такая проверка осуществляется вкратце следующим образом. Подставим $\mu^j \lambda^n \hat{u}$ вместо u_j^n в (6.85) с $\Delta t = 0$, где $\hat{u}$ — постоянный вектор; это дает возможность найти нетривиальное решение уравнения $\det[M(\lambda, \mu)] = 0$ (здесь $M = D_0 + \mu D_1 + \dots + \mu^k D_k$). Если найдутся такие корни $\mu_i(\lambda)$, что $|\mu_i(\lambda)| = 1$ и $|\lambda| > 1$, то уравнения будут неустойчивыми, потому что не выполняется условие (1) основной теоремы. С другой стороны, если имеется корень $\mu_0(\lambda)$ кратности k с $|\mu_0(\lambda)| < 1$, то найдем те векторы u , для которых $M(\lambda, \mu_0(\lambda))u = 0$. Если среди них нет k линейно независимых, то дополним их векторами u , для которых $M^2 u = 0$, и т. д., пока не получим k линейно независимых векторов. Они образуют подпространство $\mathcal{R}(\mu_0(\lambda))$.

Этот процесс повторяется для всех корней $\mu_i(\lambda)$ с $|\mu_i(\lambda)| < 1$, и таким образом строится все подпространство $\mathcal{R}$. Далее смотрим, совпадает ли размерность этого подпространства с числом граничных условий на левой границе; если это так, то применяем эти условия к произвольной линейной комбинации векторов, образующих базис в $\mathcal{R}$. Условие существования нетривиального решения соответствующей системы уравнений дает уравнение относительно λ . У этого последнего не должно быть решений с $|\lambda| > 1$. Если все это выполнено для обеих границ, то задача удовлетворяет критерию Годунова — Рябенского. Отметим, что хотя подпространство $\mathcal{R}$ и зависит от λ , его размерность для устойчивой системы не зависит от λ при $|\lambda| > 1$. В самом деле, корни $\mu_i(\lambda)$ непрерывно зависят от λ , а размерность $\mathcal{R}$ может изменяться лишь при переходе некоторого μ_i из области $|\mu_i| < 1$ в область $|\mu_i| > 1$; но такой переход невозможен, потому что $|\mu_i|$ никогда не равняется единице при $|\lambda| > 1$.

Из приведенных выше рассуждений о хорошей обусловленности можно теперь видеть, какова структура спектра семейства разностных операторов $\{C_J\}$. Он состоит из трех частей, порождаемых соответственно основными разностными уравнениями (6.85) и граничными условиями (6.86) и (6.87) слева и

справа и проявляющихся в нарушении соответственно условий (1), (2a) и (2b) основной теоремы. Первая часть состоит из корней λ уравнения $\det[M(\lambda, \mu)] = 0$ с $|\mu| = 1$ и совпадает со спектром семейства (6.85), где j пробегает значения от $-\infty$ до $+\infty$ и граничные условия отсутствуют. Таким образом, эти точки λ совпадают с собственными значениями соответствующих матриц перехода (при $\Delta t = 0$) и образуют p замкнутых кривых в комплексной плоскости, когда μ пробегает (не более p раз) единичную окружность. Вторую часть спектра образуют такие значения λ , для которых разностные уравнения (6.85) при $j = 0, 1, 2, \dots$ совместно с граничными условиями (6.86) при $j = 0$ имеют решение вида $u^n = \lambda^n u$ при ненулевом граничном значении u . Это возможно в случае, когда размерность $\mathcal{R}$ не равняется числу строк системы (6.86); рассматривая различные связные области, на которые комплексная плоскость разбивается кривыми, составляющими первую часть спектра, мы видим, что если это неравенство имеет место в какой-нибудь одной точке такой области, то оно выполняется и всюду в этой области, и вся эта область добавляется к спектру семейства. Например, для задачи, приведенной в начале § 6.6, при дополнении ее левым граничным условием к спектру добавится вся область, лежащая внутри окружности $|\lambda - (1 - r)| = r$. Кроме того, к спектру может добавиться еще ряд точек, если граничный оператор $(\lambda\beta_{00} + \alpha_{00}) + \dots + (\lambda\beta_{0k_0} + \alpha_{0k_0})$, соответствующий условию (6.86), переводит базис подпространства $\mathcal{R}$ в систему линейно зависимых векторов. Такая потеря размерности имеет место либо тождественно по λ — и тогда разностная схема не может быть устойчивой, — либо не более чем в конечном числе точек. Наконец, третья часть спектра получается при аналогичном рассмотрении правого граничного условия.

Теперь можно сказать несколько слов о том, что изменяется, когда матричные коэффициенты A_k и B_k являются кусочно непрерывными функциями от x (для детального ознакомления с этим вопросом мы отсылаем читателя к статьям Годунова и Рябенского). В спектр будут входить:

1) в каждой точке x отрезка $0 \leq x \leq 1$ те корни уравнения $\det[M(x, \lambda, \mu)] = 0$, для которых $|\mu| = 1$, причем в точках разрыва нужно рассматривать как $M(x + 0, \lambda, \mu)$, так и $M(x - 0, \lambda, \mu)$;

2) точки, возникающие при учете граничных условий;

3) на разрывах те точки λ , для которых $\dim \mathcal{R}(x + 0, \lambda) + \dim \mathcal{S}(x - 0, \lambda) \neq p$ или эти два подпространства не являются линейно независимыми

Критерий Годунова — Рябенского будет выполнен, если полученный таким образом полный спектр будет лежать в замкнутом единичном круге.

§ 6.8. Заключение

В настоящее время теория смешанных краевых задач еще не получила такого широкого развития, как теория однородных краевых задач. Причина такого различия заключается в том, что использовать методы Фурье для исследования уравнений с постоянными коэффициентами просто в одном случае и сложно в другом. Тем не менее последние достижения в этой области можно считать фрагментами той теории, которую еще предстоит построить.

Энергетический метод является мощным инструментом для исследования конкретных уравнений или конкретных классов уравнений. Хотя его применение подчас является довольно сложным, он дает возможность эффективно учитывать граничные условия и переменность коэффициентов. Кроме того, его можно использовать не только для доказательства устойчивости данного разностного метода, но также и для правильного выбора самого метода. Однако энергетический метод позволяет получать только достаточные условия устойчивости, которые могут сильно отличаться от необходимых (особенно в том случае, когда учитываются граничные условия). Простой пример волнового уравнения дает некоторое представление о том, как за счет подходящего выбора энергии можно сократить этот разрыв.

Дополнением энергетического метода служит спектральная теория Годунова — Рябенского. Хотя ее применение в любом практическом случае является исключительно сложным, она имеет совершенно общий характер и хорошо помогает понять внутреннюю сущность неустойчивости. Но с помощью этой теории получаются только необходимые условия устойчивости, которые, к сожалению, также могут быть слишком слабыми, как это показывает сравнение с условием фон Неймана для краевых задач с периодическими решениями. С другой стороны, Крайс [1965] показал, что для нашей простой модельной задачи (6.22) эти условия достаточны для устойчивости. Больше того, при несколько более сильных предположениях он распространил затем этот результат на системы гиперболических уравнений.

Существует еще один метод для исследования смешанных задач, который мы пока не рассматривали, а именно метод Винера — Хопфа. Этот метод вслед за Крейном [1958] использо-

вал Стрэнг [1964b]. Его достоинством является то, что он применим к неявным разностным схемам. Рассмотрим на всей действительной прямой систему скалярных уравнений с постоянными коэффициентами

$$\sum_{-K}^K {}_{(k)} b_k u_{j+k}^{n+1} = \sum_{-K}^K {}_{(k)} a_k u_{j+k}^n, \quad (6.99)$$

которую мы запишем в виде

$$Bu^{n+1} = Au^n. \quad (6.100)$$

Сужение этих операторов A и B на положительную полуось, для которого $u_j^n = 0$ при $j \leq 0$, обозначим через A_+ и B_+ . Тогда устойчивость смешанной задачи на этой полуоси определяется тем, будет ли $\|(B_+^{-1}A_+)^n\| \leq \text{const}$ при $n\Delta t \leq 1$. Если обозначить $\sum_k a_k e^{ik\theta}$ через $a(\theta)$, $\sum_k b_k e^{ik\theta}$ через $b(\theta)$, а $a(\theta)/b(\theta)$ через $r(\theta)$, то метод Фурье показывает, что $\|(\bar{B}^{-1}A)\| = \sup_{\theta} |r(\theta)|$; поэтому условием фон Неймана будет неравенство $|r(\theta)| \leq 1$. Вот теорема, доказанная Стрэнгом¹⁾.

Тогда и только тогда оператор B_+ обратим и имеет место устойчивость, когда

(1) $b(\theta) \neq 0$ для действительных θ ,

(2) индекс $(b) \equiv (2\pi)^{-1} \int_{-\pi}^{\pi} d[\arg b(\theta)] = 0$,

(3) $|r(\theta)| \leq 1$ для действительных θ .

Условия (1) и (2) являются условиями того, что дискретный оператор Винера — Хопфа B_+ обратим; они эквивалентны предположению о том, что полином $z^K \beta(z) = \sum_k b_k z^{K+k}$ имеет в точности K нулей в открытом круге $|z| < 1$ и K в его внешности $|z| > 1$. Тогда этот полином может быть разложен на множители следующим образом:

$$\beta(z) = s(z) t(z) = \left[\sum_0^K {}_{(k)} s_k z^k \right] \left[\sum_{-K}^0 {}_{(k)} t_k z^k \right], \quad s_0 = 1, \quad (6.101)$$

где $s \neq 0$ при $|z| \leq 1$ и $t \neq 0$ при $|z| \geq 1$; s и t в отдельности также удовлетворяют условиям (1) и (2). Это разложение играет основную роль при доказательстве теоремы. В самом деле,

¹⁾ Стрэнг рассматривает также тот случай, когда коэффициенты a_k и b_k непрерывно зависят от x и Δx .

точно так же, как полиномы a и b отвечают определенным на всей оси матричным операторам A и B с элементами $A_{ij} = a_{j-i}$ и $B_{ij} = b_{j-i}$, так и полиномы s и t отвечают матрицам, имеющим соответственно верхнюю и нижнюю треугольную форму. Таким образом, $B = ST$, $B_+ = S_+T_+$ и

$$(B_+^{-1}A_+)^n = T_+^{-1}(S_+^{-1}A_+T_+^{-1})^n T_+.$$

Далее, так как матрицы S_+^{-1} и T_+^{-1} имеют треугольную форму, то $S_+^{-1}A_+T_+^{-1} = (S^{-1}AT^{-1})_+^{-1}$, а для определенных на всей оси операторов имеем $S^{-1}AT^{-1} = S^{-1}T^{-1}A = B^{-1}A$. Следовательно,

$$(B_+^{-1}A_+)^n = T_+^{-1}(B^{-1}A)_+^n T_+,$$

так что устойчивость определяется условием фон Неймана (3).

Отметим, что в этом случае условия теоремы Годунова — Рябенского на границе $x = 0$ выполняются тривиальным образом. Действительно, число нулей функции $z^k[\lambda\beta(z) - \alpha(z)]$ в единичном круге одно и то же для всех $|\lambda| > 1$; значит, оно такое же, как и для самого β , т. е. равно K . Далее, поскольку в задаче на полуоси, для которой (6.99) выполняется при $j = 1, 2, \dots$, никаких граничных условий справа не налагается, то должно быть $b_k = 0$ при $k \geq 1$, и можно считать, что $b_{-k} \neq 0$. Таким образом, нужно предположить, что на каждом временном шаге выполняются условия $u_0 = u_{-1} = \dots = u_{-K+1} = 0$, которые препятствуют распространению любого из K затухающих колебаний внутрь рассматриваемой области.

В этой главе мы сосредоточили внимание на задачах с одним пространственным переменным. Большинство методов и результатов можно легко распространить на задачи с многими переменными, если рассматриваемая область по-прежнему является полупространством²⁾; однако рассмотрение более сложных границ, которые могут представиться в многомерном случае, намного более трудно, и до сих пор мало что сделано в этом направлении.

¹⁾ Действительно, пусть

$$U = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad U_+ = \begin{bmatrix} 0 & 0 \\ 0 & d \end{bmatrix}, \quad V = \begin{bmatrix} e & f \\ g & h \end{bmatrix}, \quad V_+ = \begin{bmatrix} 0 & 0 \\ 0 & h \end{bmatrix}.$$

Тогда матрица

$$(UV)_+ = \begin{bmatrix} 0 & 0 \\ 0 & cf + dh \end{bmatrix}$$

равняется U_+V_+ , если $c = 0$ или $f = 0$, т. е. если U имеет верхнюю или V имеет нижнюю треугольную форму.

²⁾ С этой задачей для дифференциальных уравнений читатель может познакомиться по работе Херша [1963].

Глава 7

МНОГОСЛОЙНЫЕ РАЗНОСТНЫЕ УРАВНЕНИЯ

§ 7.1. Обозначения

Как и для обыкновенных дифференциальных уравнений, при построении разностных уравнений повышенной точности для уравнений с частными производными число используемых временных слоев часто превышает минимальное их число, требуемое дифференциальными уравнениями. Например, улучшение полученного ранее простого двухслойного неявного уравнения для задачи теплопроводности

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \sigma \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{(\Delta x)^2} + O(\Delta t) + O(\Delta x^2)$$

получается переходом к трехслойному неявному уравнению

$$\begin{aligned} \frac{\frac{3}{2} u_j^{n+1} - 2u_j^n + \frac{1}{2} u_j^{n-1}}{\Delta t} &= \\ &= \sigma \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{(\Delta x)^2} + O(\Delta t^2) + O(\Delta x^2). \end{aligned} \quad (7.1)$$

В случае когда применяется это уравнение, нужны начальные данные для двух значений времени, скажем для $t = 0$ и $t = \Delta t$ ($n = 0$ и 1), чтобы можно было начать вычисления. Эти данные могут быть получены с помощью более простого уравнения с, быть может, более мелкой сеткой, или разложением в степенные ряды, или каким-либо другим способом; в любом случае, если эти данные уже имеются, то решение шаг за шагом можно построить по уравнению (7.1) и соответствующим граничным условиям.

Более общим образом, предположим, что краевая задача

$$\frac{\partial u}{\partial t} = Au, \quad u = u(t) \quad (7.2)$$

аппроксимируется с помощью $(q + 1)$ -слойной формулы

$$B_q u^{n+q} + B_{q-1} u^{n+q-1} + \dots + B_0 u^n = 0, \quad (7.3)$$

где $B_q, \dots, B_0$ обозначают конечноразностные операторы, а u^n — некоторая аппроксимация $u(n\Delta t)$. Предположим, что уравнение (7.3) вместе с надлежащими граничными условиями

единственным образом разрешимо относительно u^{n+q} и что это решение непрерывно зависит от $u^{n+q-1}, \dots, u^n$. Это значит, что уравнение (7.3) эквивалентно уравнению

$$u^{n+q} = C_{q-1}u^{n+q-1} + \dots + C_0u^n, \quad (7.4)$$

где

$$C_j = C_j(\Delta t) = -B_q^{-1}B_j, \quad j=0, \dots, q-1,$$

суть линейные ограниченные операторы. Как и в предыдущих рассмотрениях, мы предполагаем, что приращения пространственных переменных $\Delta x, \Delta y$ и т. д. являются заданными функциями от Δt .

§ 7.2. Вспомогательное банахово пространство

Будем рассматривать q -мерный вектор-столбец, компоненты которого являются элементами пространства $\mathcal{B}$, как элемент вспомогательного банахова пространства $\tilde{\mathcal{B}}$. Норму в нем можно определить по-разному, например так:

$$\left\| \begin{bmatrix} u \\ v \\ \vdots \end{bmatrix} \right\| = (\|u\|^2 + \|v\|^2 + \dots)^{1/2}.$$

Если положить

$$\tilde{\varphi}^n = \begin{bmatrix} u^{n+q-1} \\ \vdots \\ u^n \end{bmatrix}$$

и

$$\tilde{C} = \begin{bmatrix} C_{q-1} & C_{q-2} & \dots & C_1 & C_0 \\ I & 0 & \dots & 0 & 0 \\ 0 & I & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & I & 0 \end{bmatrix} = \tilde{C}(\Delta t), \quad (7.5)$$

где I — тождественный оператор, то уравнение (7.4) запишется в виде

$$\tilde{\varphi}^{n+1} = \tilde{C}(\Delta t) \tilde{\varphi}^n. \quad (7.6a)$$

Это равносильно введению новых зависимых переменных; именно, если обозначить u^{n-1} через v^n , u^{n-2} через w^n и т. д., то

уравнение (7.4) можно заменить системой

$$\begin{aligned} u^{n+1} &= C_{q-1}u^n + C_{q-2}v^n + C_{q-3}w^n + \dots, \\ v^{n+1} &= u^n, \\ w^{n+1} &= v^n \end{aligned} \quad (7.6b)$$

и т. д.

Таким образом, $(q+1)$ -слойную задачу можно свести к двух-слойной задаче (ибо верхние индексы, входящие в уравнения системы, принимают лишь значения n и $n+1$).

Исходные начальные данные предполагаются заданными при $t=0$; значения u при $t=\Delta t, 2\Delta t, \dots, (q-1)\Delta t$ должны быть определены по исходным начальным данным с помощью некоторой аппроксимации, что даст возможность начать вычисления с помощью уравнения (7.3). Предположим сначала, что эти начальные данные вычислены точно (это требование позднее будет ослаблено), т. е. что $\tilde{\Phi}^0$ задается равенством

$$\tilde{\Phi}^0 = \begin{bmatrix} E((q-1)\Delta t)u^0 \\ E((q-2)\Delta t)u^0 \\ \vdots \\ u^0 \end{bmatrix},$$

где исходные начальные данные определяются элементом u^0 (пространства $\mathcal{B}$), а $E(t)$ — разрешающий оператор для уравнения (7.2). Тогда можно написать

$$\tilde{\Phi}^0 = \tilde{S}\tilde{u}^0,$$

где

$$\tilde{u}^0 = \begin{bmatrix} u^0 \\ u^0 \\ \vdots \\ u^0 \end{bmatrix},$$

а $\tilde{S}$ — диагональная матрица

$$\tilde{S} = \begin{bmatrix} E((q-1)\Delta t) & & & & \\ & 0 & & & \\ & & E((q-2)\Delta t) & & \\ & & & \ddots & \\ 0 & & & & I \end{bmatrix}. \quad (7.7)$$

Таким образом, приближенное решение краевой задачи (7.2) дается формулой $\tilde{\varphi}^n = \tilde{C}(\Delta t)^n \tilde{S} \tilde{u}^0$. Мы ожидаем, что это — аппроксимация выражения

$$\begin{bmatrix} u((n+q-1)\Delta t) \\ u((n+q-2)\Delta t) \\ \vdots \\ u(n\Delta t) \end{bmatrix} = E(n\Delta t) \tilde{S} \tilde{u}^0;$$

поэтому нам интересно сравнить $\tilde{C}(\Delta t)^n \tilde{S} \tilde{u}^0$ с $E(n\Delta t) \tilde{S} \tilde{u}^0$.

В пределе при $\Delta t \rightarrow 0$ мы придем к элементам пространства $\tilde{\mathcal{B}}$, все q компонент которых равны между собой, а именно получим

$$\begin{bmatrix} u(t) \\ \vdots \\ u(t) \end{bmatrix} = E(t) \begin{bmatrix} u^0 \\ \vdots \\ u^0 \end{bmatrix},$$

что запишем в виде $\tilde{u}(t) = E(t) \tilde{u}^0$. К этому пределу стремится также $E(t) \tilde{S} \tilde{u}^0$, ибо $\tilde{S} \rightarrow I$ при $\Delta t \rightarrow 0$. Удобно обозначить через $\tilde{\mathcal{B}}^0$ подпространство пространства $\tilde{\mathcal{B}}$, состоящее из элементов, имеющих равные компоненты; другими словами, элемент

$$\tilde{\varphi} = \begin{bmatrix} u \\ v \\ w \\ \vdots \\ \cdot \end{bmatrix}$$

пространства $\tilde{\mathcal{B}}$ принадлежит подпространству $\tilde{\mathcal{B}}^0$ тогда и только тогда, когда $u = v = w = \dots$. Элементы $\tilde{u}(t)$ и $\tilde{u}^0$ лежат в подпространстве $\tilde{\mathcal{B}}^0$, но их аппроксимации — нет, они лежат лишь поблизости от него.

На практике начальные значения получают не с помощью точного вычисления $\tilde{\varphi}^0 = \tilde{S} \tilde{u}^0$, а с помощью аппроксимации $\tilde{\varphi}^0 = \tilde{\varphi}$. Мы будем требовать, чтобы приближенное решение сходилось к $\tilde{u}(t)$, когда $\Delta t \rightarrow 0$ и $\tilde{\varphi} \rightarrow \tilde{u}^0$ независимо друг от друга; мы не требуем, чтобы $\tilde{\varphi}$ лежало в подпространстве $\tilde{\mathcal{B}}^0$, а только чтобы его предел принадлежал этому подпространству.

§ 7.3. Теорема об эквивалентности

Многослойная разностная формула (7.6a) называется *устойчивой*, если для некоторого положительного τ операторы

$$\tilde{C}(\Delta t)^n \text{ при } \begin{cases} 0 < \Delta t < \tau, \\ 0 \leq n \Delta t \leq T \end{cases} \quad (7.8)$$

равномерно ограничены. Иными словами, эквивалентная двухслойная система (7.6b) должна быть устойчивой в смысле определения, данного в гл. 3.

Условие согласованности состоит в том, что существует плотное в $\mathcal{B}$ множество U^0 возможных начальных элементов для точных решений, такое, что если $u^0 \in U^0$ и $\varepsilon > 0$ и если положить

$$\tilde{u}(t) = \begin{bmatrix} E(t) u^0 \\ \vdots \\ E(t) u^0 \end{bmatrix},$$

то

$$\|(\tilde{C} \Delta t) - E(\Delta t)\tilde{S}\tilde{u}(t)\| < \varepsilon \Delta t \quad (7.9)$$

для достаточно малых Δt и $0 \leq t \leq T$.

Это не что иное, как двойственная формулировка согласованности, данная для двухслойных уравнений в (3.6b). Все $\tilde{u}(t)$, фигурирующие в настоящем условии, лежат в $\tilde{\mathcal{B}}^0$.

Аппроксимация называется *сходящейся*, если для любых последовательностей $\Delta_j t$ и n_j , таких, что $\Delta_j t > 0$, $\Delta_j t \rightarrow 0$ и $n_j \Delta_j t \rightarrow t$ при $j \rightarrow \infty$ ($0 \leq t \leq T$), и любого $\tilde{u} \in \tilde{\mathcal{B}}^0$ будет

$$\lim_{\substack{j \rightarrow \infty \\ \tilde{\psi} \rightarrow \tilde{u}}} \|\tilde{C}(\Delta_j t)^{n_j} \tilde{\psi} - E(t)\tilde{u}\| = 0. \quad (7.10)$$

Теорема об эквивалентности гласит, как и ранее, что

Для корректно поставленной краевой задачи и аппроксимации, удовлетворяющей условию согласованности, устойчивость необходима и достаточна для сходимости.

Доказательство аналогично приведенному в гл. 3; единственное существенное различие связано с различием между пространствами $\mathcal{B}$ и $\mathcal{B}^0$. Для того чтобы доказать, что из сходимости следует устойчивость, допустим, что аппроксимация сходится, и пусть $\tilde{\psi}$ — некоторый элемент пространства $\mathcal{B}$. Мы утверждаем, что величины

$$\|\tilde{C}(\Delta t)^n \tilde{\psi}\| \text{ при } \begin{cases} 0 \leq \Delta t < \tau, \\ 0 \leq n \Delta t \leq T \end{cases}$$

ограничены; действительно, в противном случае мы выбрали бы из них такую подпоследовательность, что

$$\|\tilde{C}(\Delta_j t)^{n_j} \tilde{\psi}\| \rightarrow \infty \quad \text{при } j \rightarrow \infty,$$

а затем из нее выбрали бы такую подпоследовательность, для которой $n_j \Delta_j t$ имеет предел, скажем t ($0 \leq t \leq T$) при $j \rightarrow \infty$. Тогда элементы

$$\tilde{\psi}_j = \frac{\tilde{\psi}}{\|\tilde{C}(\Delta_j t)^{n_j} \tilde{\psi}\|^{1/2}}$$

стремились бы к нулевому элементу

$$\tilde{u}^0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix},$$

лежащему в подпространстве $\tilde{\mathcal{B}}^0$, и мы имели бы

$$\|\tilde{C}(\Delta_j t)^{n_j} \tilde{\psi}_j\| \rightarrow \infty \quad \text{при } j \rightarrow \infty,$$

в то время как в силу сходимости $\tilde{\psi}_j$ к $\tilde{u}^0$ мы должны были бы иметь

$$\|\tilde{C}(\Delta_j t)^{n_j} \tilde{\psi}_j\| \rightarrow \|E(t) \tilde{u}^0\| = 0 \quad \text{при } j \rightarrow \infty.$$

Поэтому существует такая граница $K_1(\tilde{\psi})$, что

$$\|\tilde{C}^n(\Delta t)^n \psi\| < K_1(\tilde{\psi}) \quad \text{для} \quad \begin{cases} 0 < \Delta t < \tau, \\ 0 \leq n \Delta t \leq T, \\ \tilde{\psi} \in \tilde{\mathcal{B}}, \end{cases}$$

и устойчивость следует из принципа равномерной ограниченности.

Для доказательства того, что из устойчивости в свою очередь следует сходимость, возьмем

$$\Delta_j t > 0, \quad \lim_{j \rightarrow \infty} \Delta_j t = 0, \quad \lim_{j \rightarrow \infty} n_j \Delta_j t = t, \quad 0 \leq t \leq T,$$

и введем обозначение

$$\tilde{\chi}_j = [\tilde{C}(\Delta_j t)^{n_j} - E(n_j \Delta_j t)] \tilde{S} \tilde{u}^0,$$

где u^0 принадлежит множеству U^0 , входящему в определение согласованности. Тогда $\tilde{u}(t) = E(t) \tilde{u}^0$ является точным решением и

$$\tilde{\chi}_j = \sum_{k=0}^{n_j-1} \tilde{C}(\Delta_j t)^k [\tilde{C}(\Delta_j t) - E(\Delta_j t)] E((n_j - k - 1) \Delta_j t) \tilde{S} \tilde{u}^0.$$

Далее, $\tilde{S}$ перестановочно с $E(t)$ (см. равенство (7.7)); следовательно, в соответствии с условием согласованности (7.9) при $\varepsilon > 0$

$$\| [\tilde{C}(\Delta_j t) - E(\Delta_j t)] E((n_j - k - 1)\Delta_j t) \tilde{S}\tilde{u}^0 \| < \varepsilon \Delta_j t$$

для достаточно малых положительных $\Delta_j t$. Если K_2 — общая граница операторов (7.8), то, применяя неравенство треугольника, получаем, что для достаточно малых $\Delta_j t > 0$

$$\| \tilde{\chi}_j \| \leq \sum_{k=0}^{n_j-1} K_2 \varepsilon \Delta_j t = K_2 \varepsilon n_j \Delta_j t \leq K_2 \varepsilon T.$$

Значит, в силу произвольности ε , $\tilde{\chi}_j \rightarrow 0$ при $j \rightarrow \infty$. Как и в гл. 3, можно теперь показать, что

$$[\tilde{C}(\Delta_j t)^{n_j} - E(t)] \tilde{S}\tilde{u}^0 \rightarrow 0,$$

где t является пределом $n_j \Delta_j t$. Поэтому (опять как в гл. 3), если u — произвольный элемент пространства $\mathcal{B}$, то мы можем найти такую последовательность $\{u_m\}$ элементов множества U^0 , что $u_m \rightarrow u$ при $m \rightarrow \infty$, и если положить

$$\tilde{u}_m = \begin{bmatrix} u_m \\ \vdots \\ u_m \end{bmatrix}, \quad \tilde{u} = \begin{bmatrix} u \\ \vdots \\ u \end{bmatrix},$$

то

$$\begin{aligned} [\tilde{C}(\Delta_j t)^{n_j} - E(t)] \tilde{S}\tilde{u} &= [\tilde{C}(\Delta_j t)^{n_j} - E(t)] \tilde{S}\tilde{u}_m + \\ &+ \tilde{C}(\Delta_j t)^{n_j} \tilde{S}(\tilde{u} - \tilde{u}_m) - E(t) \tilde{S}(\tilde{u} - \tilde{u}_m). \end{aligned}$$

Последние два члена этого равенства можно сделать сколь угодно малыми за счет выбора достаточно большого m в силу равномерной ограниченности¹⁾ операторов $\tilde{C}^{n_j}(\Delta_j t)$, $E(t)$ и $\tilde{S}$; первый член в правой части можно сделать сколь угодно малым за счет выбора достаточно большого j .

Для завершения доказательства сходимости нужен еще один шаг, который не был необходим в гл. 3, а именно

$$\begin{aligned} \tilde{C}(\Delta_j t)^{n_j} \tilde{\psi} - E(t) \tilde{u} &= [\tilde{C}(\Delta_j t)^{n_j} - E(t)] \tilde{S}\tilde{u} + \\ &+ \tilde{C}(\Delta_j t)^{n_j} (\tilde{\psi} - \tilde{S}\tilde{u}) - E(t) (\tilde{u} - \tilde{S}\tilde{u}). \end{aligned}$$

¹⁾ Заметим, что оператор $\tilde{S}$ является функцией от Δt , ограниченной при $0 \leq \Delta t \leq \tau$; в самом деле, $\tilde{S} \rightarrow I$ при $\Delta t \rightarrow 0$.

При $\Delta t \rightarrow 0$ имеем $\tilde{S} \rightarrow I$; поэтому если $\tilde{\psi} \rightarrow \tilde{u}$, то $\|\tilde{\psi} - \tilde{S}\tilde{u}\| \rightarrow 0$, $\|\tilde{u} - \tilde{S}\tilde{u}\| \rightarrow 0$, и мы видим, что

$$\lim_{\substack{J \rightarrow \infty \\ \tilde{\psi} \rightarrow \tilde{u}}} \|\tilde{C}(\Delta, t)^n \tilde{\psi} - E(t) \tilde{u}\| = 0$$

для произвольного $\tilde{u}$ из $\tilde{\mathcal{H}}^0$, что и требовалось доказать.

§ 7.4. Согласованность и порядок точности

При исследовании согласованности и точности удобно иметь дело с разностными уравнениями в их простейшей форме (7.3), не разрешенной относительно u^{n+q} . Неопределенность вследствие возможности умножения слева на любую невырожденную матрицу устраняется предположением о том, что для любой достаточно гладкой функции $u(t)$ разностный оператор в левой части аппроксимирует оператор $\partial/\partial t - A$, т. е.

$$B_q u(t + q \Delta t) + \dots + B_0 u(t) - \left(\frac{\partial}{\partial t} - A \right) u(t) = o(1) \quad \text{при } \Delta t \rightarrow 0. \quad (7.11)$$

Как и в гл. 4, определим *порядок точности* как наибольшее действительное число ρ , для которого

$$\|B_q u(t + q \Delta t) + \dots + B_0 u(t)\| = O[(\Delta t)^\rho] \quad \text{при } \Delta t \rightarrow 0 \quad (7.12)$$

для всех достаточно гладких точных решений дифференциального уравнения $\partial u/\partial t = Au$. Выражение в левой части (7.12) называется погрешностью аппроксимации и может быть оценено путем разложения u в ряд Тэйлора.

Чтобы связать это с условием согласованности (7.9), которое выражается через операторы $C_j = -B_q^{-1} B_j$, мы уточним соответствующие рассуждения для краевых задач с постоянными коэффициентами и условием периодичности решения, приведенные в гл. 4. Мы покажем ниже, что $B_q^{-1} = O(\Delta t)$. Это утверждение следует из соотношения (7.11) и ограниченности операторов $C_1, C_2, \dots, C_{q-1}$, но доказать его труднее, чем в случае двухслойных уравнений. Следовательно, *если разностная схема устойчива и погрешность аппроксимации стремится к нулю при $\Delta t \rightarrow 0$, то условие согласованности (7.9) выполняется.* Только первая компонента вектора $[\tilde{C}(\Delta t) - E(\Delta t)]\tilde{S}\tilde{u}(t)$ отлична от нуля, так что нам нужно лишь показать, что при любом $\varepsilon > 0$

$$\|[C_{q-1}E((q-1)\Delta t) + \dots + C_0 - E(q\Delta t)]u(t)\| < \varepsilon \Delta t$$

для любого элемента $u(0)$ из некоторого множества U^0 , плотного в $\mathcal{H}$, и для всех достаточно малых Δt .

Как и в гл. 4, банахово пространство $\mathcal{B}$ есть гильбертово пространство функций $u(x)$, удовлетворяющих условию периодичности, причем за норму принята гильбертова норма. Пусть $\mathcal{B}_m$ — подпространство тригонометрических полиномов степени m . Если u_0 представляет функцию $u_0(x)$ из подпространства $\mathcal{B}_m$ и если в равенстве (7.11) положить $u(t) = tu_0$, то

$$Lu_0 = o(1) \quad \text{при} \quad \Delta t \rightarrow 0, \quad (7.13)$$

где

$$L = \Delta t (qB_q + (q-1)B_{q-1} + \dots + B_1) - I. \quad (7.14)$$

Так как $\mathcal{B}_m$ имеет конечный базис, каждый элемент которого удовлетворяет соотношению (7.13), то $\|L\|_m < \varepsilon_1$ для достаточно малых Δt , где ε_1 — произвольно малое положительное число, $\|L\|_m$ — норма оператора L в подпространстве $\mathcal{B}_m$. Поэтому оператор $I + L$ имеет обратный в этом подпространстве, а именно

$$(I + L)^{-1} = I + M,$$

где

$$\|M\|_m < \varepsilon_2 = \frac{\varepsilon_1}{1 - \varepsilon_1};$$

таким образом,

$$[qB_q + \dots + B_1]^{-1} = \Delta t (I + M).$$

Из определения $C_1 = -B_q^{-1}B_1$ следует, что

$$B_q^{-1} = [qI - (q-1)C_{q-1} - \dots - C_1] \Delta t (I + M). \quad (7.15)$$

Для рассматриваемой задачи решение $u(t)$ лежит в $\mathcal{B}_m$, если $u(0)$ лежит в $\mathcal{B}_m$, потому что A в уравнении (7.2) является дифференциальным оператором с постоянными коэффициентами. Кроме того, для этого решения будет $\partial u / \partial t - Au = 0$, так что в силу (7.11)

$$\| [B_q E(q \Delta t) + B_{q-1} E((q-1) \Delta t) + \dots + B_0] u(t) \| < \varepsilon_0 \quad (7.16)$$

для любого $\varepsilon_0 > 0$ и для всех достаточно малых Δt .

Умножая (7.15) на (7.16), получаем

$$\| [E(q \Delta t) - C_{q-1} E((q-1) \Delta t) - \dots - C_0] u(t) \| < \varepsilon_3 \Delta t, \quad (7.17)$$

где

$$\varepsilon_3 = \left\{ \max_{0 < \Delta t < \tau} \| qI - (q-1)C_{q-1} - \dots - C_1 \| \right\} \varepsilon_0 \| I + M \|_m.$$

Входящий в последнее равенство максимум существует, поскольку мы предположили, что разностные уравнения устойчивы, откуда вытекает, что операторы $C_{q-1}(\Delta t), \dots, C_1(\Delta t)$ равномерно ограничены при $0 < \Delta t < \tau$. Неравенство вида (7.17)

имеет место для любого решения $u(t)$, принадлежащего $\mathcal{B}_m$ (это, конечно, точное решение). Если принять во внимание все подпространства $\mathcal{B}_m$, то возможные начальные элементы $u(0)$ для этих решений образуют плотное множество в $\mathcal{B}$, чем согласованность и доказана.

§ 7.5. Пример Дюфорта и Франкела

Чтобы проиллюстрировать сказанное выше, исследуем согласованность предложенной Дюфортом и Франкелом [1953] разностной схемы

$$\frac{u_j^{n+1} - u_j^{n-1}}{2 \Delta t} = \sigma \frac{u_{j+1}^n - u_j^{n+1} - u_j^{n-1} + u_{j-1}^n}{(\Delta x)^2} \quad (7.18)$$

для численного решения уравнения диффузии

$$\frac{\partial u}{\partial t} = \sigma \frac{\partial^2 u}{\partial x^2}, \quad \sigma = \text{const} > 0. \quad (7.19)$$

Как будет показано ниже, уравнение (7.18) всегда устойчиво. Если члены уравнения (7.18) заменить рядами Тэйлора

$$u_j^{n+1} = \left(u + \Delta t \frac{\partial u}{\partial t} + \dots \right)_j^n \text{ и т. д.}$$

и сравнить результат с (7.19), то мы получим погрешность аппроксимации порядка $(\Delta t / \Delta x)^2$. Точнее говоря, для каждой достаточное число раз дифференцируемой функции $u(x, t)$

$$\begin{aligned} \frac{u_j^{n+1} - u_j^{n-1}}{2 \Delta t} - \sigma \frac{u_{j+1}^n - u_j^{n+1} - u_j^{n-1} + u_{j-1}^n}{(\Delta x)^2} - \\ - \left(\frac{\partial u}{\partial t} - \sigma \frac{\partial^2 u}{\partial x^2} \right)_j^n = \sigma \left(\frac{\Delta t}{\Delta x} \right)^2 \left(\frac{\partial^2 u}{\partial t^2} \right)_j^n + \\ + O[(\Delta t)^2] + O[(\Delta x)^2] + O\left(\frac{(\Delta t)^4}{(\Delta x)^2} \right). \end{aligned}$$

Поэтому для согласованности требуется, чтобы $\Delta t / \Delta x \rightarrow 0$ при $\Delta t \rightarrow 0$. Таким образом,

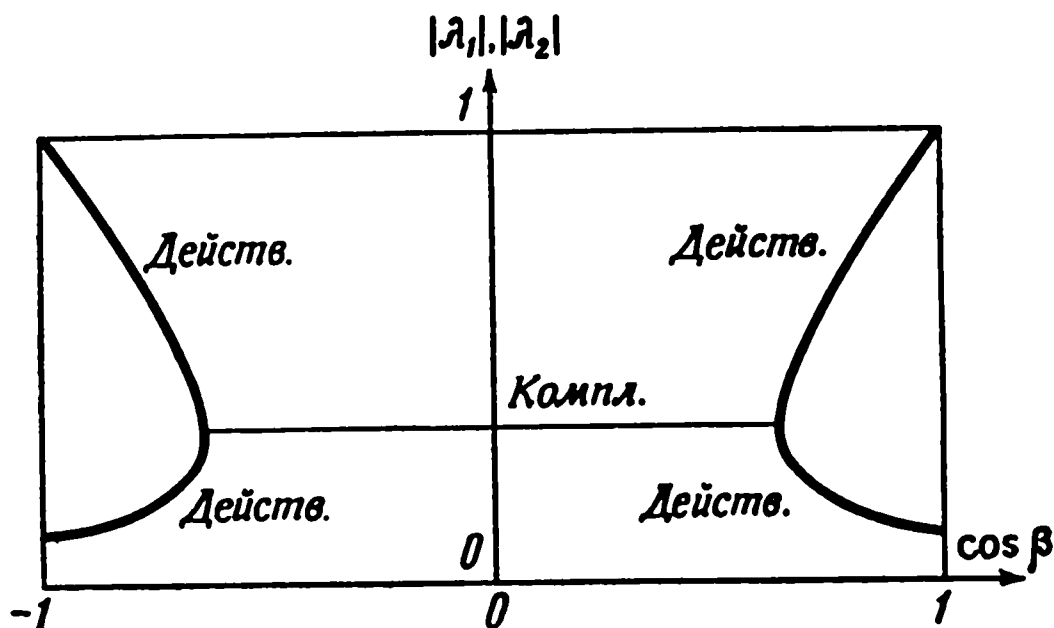
Уравнение (7.18) согласовано с уравнением (7.19) тогда и только тогда, когда Δt стремится к нулю быстрее, чем Δx ; если же $\Delta t / \Delta x$ фиксировано, скажем равно β , то уравнение (7.18) согласовано не с уравнением диффузии (7.19), а с гиперболическим уравнением

$$\frac{\partial u}{\partial t} - \sigma \frac{\partial^2 u}{\partial x^2} + \sigma \beta^2 \frac{\partial^2 u}{\partial t^2} = 0.$$

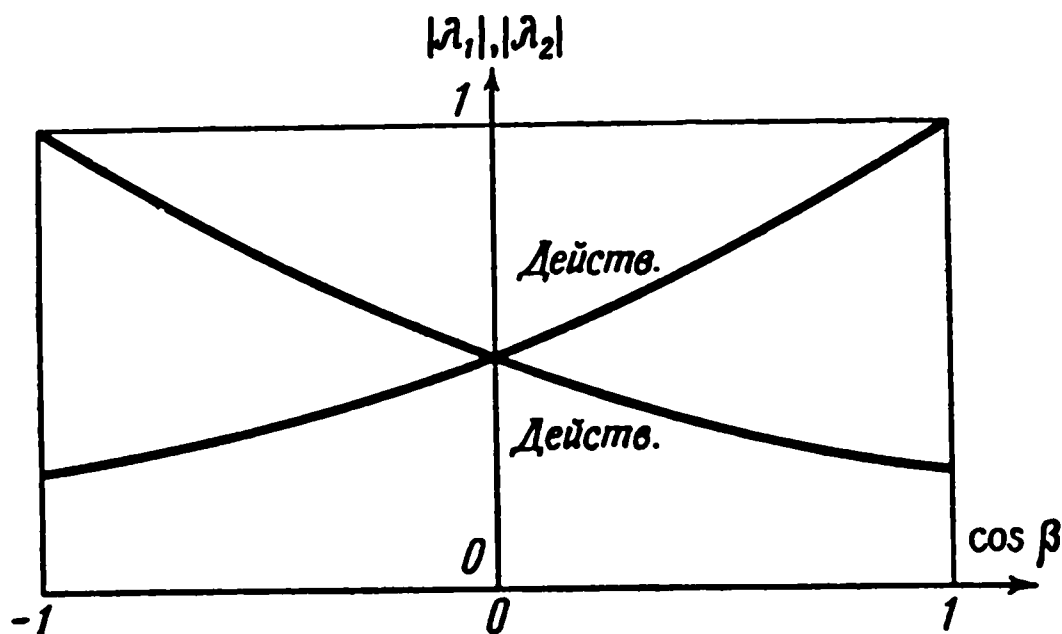
Можно показать, что разностное уравнение (7.18) всегда устойчиво. Для этого запишем эквивалентную двухслойную систему

$$\frac{u_j^{n+1} - v_j^n}{2\Delta t} = \sigma \frac{u_{j+1}^n - u_j^{n+1} - v_j^n + u_{j-1}^n}{(\Delta x)^2},$$

$$v_j^{n+1} = u_j^n.$$



Р и с. 7.1. Собственные значения матрицы перехода (7.20) разностного уравнения Дюфорта — Франкела для задачи диффузии при $\alpha > 1$.



Р и с. 7.2. То же, что и на рис. 7.1, но при $\alpha < 1$

Матрица перехода для этой системы имеет вид

$$G(\Delta t, k) = \begin{bmatrix} \frac{2\alpha}{1+\alpha} \cos \beta & \frac{1-\alpha}{1+\alpha} \\ 1 & 0 \end{bmatrix}, \quad (7.20)$$

где $\alpha = 2\sigma\Delta t/(\Delta x)^2$, $\beta = k\Delta x$; ее собственные значения равны

$$\lambda = \frac{\alpha \cos \beta \pm \sqrt{1 - \alpha^2 \sin^2 \beta}}{1 + \alpha}.$$

Зависимость модулей этих собственных значений от $\cos \beta$ для типичных случаев $\alpha > 1$ и $\alpha < 1$ схематически изображена на рис. 7.1 и 7.2 соответственно. Условие 2 из § 4.11, достаточное

для устойчивости, здесь выполняется, так как одно собственное значение удовлетворяет неравенству

$$|\lambda_1| \leq 1,$$

а другое — неравенству

$$|\lambda_2| \leq \sqrt{|1 - \alpha|/(1 + \alpha)} < 1.$$

Поэтому система Дюфорта — Франкела устойчива при любых значениях α .

§ 7.6. Резюме

Для корректно поставленных задач с постоянными коэффициентами условия сходимости для многослойных разностных уравнений практически те же, что и для двухслойных. Стремление к нулю погрешности аппроксимации при $\Delta t \rightarrow 0$ влечет за собой согласованность аппроксимации, если только разностные уравнения устойчивы ¹⁾.

Для исследования устойчивости просто пишут эквивалентную двухслойную схему, вводя новые зависимые переменные, а затем исследуют устойчивость этой схемы любым из описанных в гл. 4 методов. Необходимые дополнительные начальные данные могут быть получены с помощью любой подходящей аппроксимации, имеющей силу при $\Delta t \rightarrow 0$. В частности, их можно взять просто равными соответствующим начальным данным, $u^1 = u^0$, $u^2 = u^0$ и т. д., хотя чем лучше аппроксимация начальных данных, тем, вообще говоря, выше точность аппроксимации решения при данном Δt .

¹⁾ Эта оговорка не является необходимой для двухслойных формул, но, по-видимому, согласованность не представляет интереса, если уравнения неустойчивы.

ПРЕДИСЛОВИЕ К ЧАСТИ II

Эта часть книги посвящена приложению разностных методов к некоторым краевым задачам, возникающим в физике. Материал систематизируется в соответствии с сущностью физических задач, а не с типом описывающих их уравнений, потому что при детальном рассмотрении различных физических процессов требуются, вообще говоря, совершенно различные подходы к уравнениям одного и того же типа. Например, для гиперболических систем, которые получаются в теории переноса нейтронов, применяются методы, отличные от тех, которые нужны для гиперболических систем, возникающих в гидродинамике. Несмотря на это, многие характерные особенности являются общими для задач из различных физических областей, и мы вкратце остановимся здесь на некоторых из них.

Методы решения *параболических задач* обсуждаются главным образом в гл. 8 в связи с диффузией и теплопроводностью, хотя уравнения упругих колебаний в гл. 11 и уравнения, описывающие одновременное распространение звука и тепла в § 10.4, также по существу являются параболическими в том смысле, что для них скорость распространения сигнала бесконечна.

Методы решения *линейных гиперболических систем* подробно обсуждаются в гл. 9 в связи с уравнением переноса, особенно в § 9.6—9.16; мы снова встретимся с ними в гл. 10 при рассмотрении звуковых волн; уравнения колебаний продольно-напряженного стержня или струны, рассматриваемые в § 11.6, очень близки к гиперболическим.

В гл. 12 и 13 изучаются *нелинейные гиперболические системы*, получающиеся в гидродинамике. Вследствие нелинейности методы решения этих систем являются более сложными и более тесно связанными со спецификой гидродинамических задач. Мы не делаем попыток предложить какие-либо универсальные методы, применимые к общим нелинейным гиперболическим системам, отличным от тех систем, которые описывают законы сохранения.

Методы решения *систем линейных уравнений*, получающихся для неявных схем, рассматриваются в § 8.5 для простого

трехдиагонального случая и в § 11.5 для многодиагонального случая.

Неявные схемы для различных задач приводятся в § 8.2, 8.5 и 8.6 (параболические с одним пространственным переменным), в § 8.8 и 8.9 (параболические с несколькими пространственными переменными — методы чередующихся направлений и дробных шагов) в § 9.9 и 10.3 (гиперболические), в § 11.3—11.5 (упругие колебания). Как показано в § 11.4, неявная схема является особенно подходящей для задач динамической теории упругости.

Практические критерии устойчивости в случаях, для которых обычной теории недостаточно, обсуждаются в § 9.7, 10.4, 10.5, 11.6 и 12.16. Это те случаи, в которых ограничение, налагаемое на Δt обычной теорией (при заданных Δx_i), обеспечивает ограниченность роста ошибки в пределе при $\Delta t \rightarrow 0$, но несмотря на это, рост ошибки (если не наложить дополнительных ограничений) может оказаться неприемлемым для практически используемых значений Δt .

Глава 8

ДИФФУЗИЯ И ТЕПЛОПРОВОДНОСТЬ

§ 8.1. Примеры диффузии

В качестве примера диффузионного процесса мы рассмотрим распространение тепла в неподвижной изотропной среде (см. книгу Карслоу [1945]). Начальное распределение температуры известно; источники тепла, если они имеются, известны; требуется найти распределение температуры в последующие моменты времени.

Если распределение температуры в некоторый момент времени t обозначить через $T = T(\mathbf{x}, t)$, где T означает температуру, а $\mathbf{x}$ — вектор, характеризующий положение точки в пространстве, то закон распространения тепла можно сформулировать следующим образом: существует положительная скалярная величина $\kappa = \kappa(\mathbf{x}, T)$, называемая коэффициентом теплопроводности материала и такая, что для любого распределения тепла плотность потока тепла равна

$$\mathbf{F}(\mathbf{x}, t) = -\kappa(\mathbf{x}, T(\mathbf{x}, t)) \nabla T(\mathbf{x}, t),$$

что мы запишем кратко в следующем виде:

$$\mathbf{F} = -\kappa \nabla T. \quad (8.1)$$

Физический смысл векторного поля $\mathbf{F}$ состоит в том, что интеграл от нормальной составляющей вектора $\mathbf{F}$ по какой-либо поверхности равен потоку тепла через эту поверхность в единицу времени. Закон (8.1) гласит, что для заданной температуры T и заданной точки $\mathbf{x}$ вектор плотности потока тепла пропорционален градиенту температуры в этой точке и имеет противоположное направление. Коэффициент пропорциональности κ может зависеть также и от времени, если происходят химические реакции, но мы будем предполагать, что это не имеет места.

Пусть $E = E(\mathbf{x}, t)$ и $Q = Q(\mathbf{x}, t)$ означают соответственно внутреннюю энергию, отнесенную к единице объема, и мощность тепловых источников в единице объема (количество тепла, создаваемое за единицу времени в единице объема). Вклад в $\partial E / \partial t$ от перетекания тепла равен дивергенции $\mathbf{F}$, и поэтому в силу сохранения энергии мы имеем

$$\frac{\partial E}{\partial t} = \nabla \cdot \kappa \nabla T + Q.$$

Если $a = a(x, T)$ — удельная теплоемкость материала (отнесенная к единице объема), то для температуры получается уравнение в частных производных

$$\frac{\partial T}{\partial t} = \frac{1}{a} [\nabla \cdot \kappa \nabla T + Q]. \quad (8.2)$$

Это уравнение с надлежащими начальными и граничными условиями дает смешанную краевую задачу рассматриваемого в этой главе типа. Конкретные примеры с исследованием устойчивости будут приведены для весьма упрощенных вариантов этой задачи. Однако рассматриваемые при этом системы разностных уравнений на практике могут быть легко обобщены и применены к самому общему виду уравнения (8.2).

К уравнению этого типа приводят и другие явления. Например, это уравнение описывает диффузию вещества через порнищаемую среду, если считать T концентрацией вещества, κ коэффициентом диффузии, a равным единице, а Q удельной мощностью образования вещества при химической реакции, если таковое имеет место.

Следующим примером является распространение тепла в звезде, при котором энергия переносится как посредством излучения, так и посредством теплопроводности. Если среда неподвижна, то распространение тепла в звезде, за исключением очень тонкого поверхностного слоя, совершенно точно описывается уравнением (8.2) при следующей интерпретации символов:

a = Теплоемкость вещества на единицу объема $+ 4a_0T^3$,
 κ = Коэффициент теплопроводности вещества $+ 4a_0c_0T^3/3K\rho$,
 где a_0 — постоянная Стефана — Больцмана, c_0 — скорость света, ρ — плотность вещества, $K = K(\rho, T)$ — среднее Росселанда (коэффициент непрозрачности звездной материи). Граничным условием на поверхности звезды при некоторых предположениях будет $T = 0$.

§ 8.2. Простейшая задача теплопроводности

Полагая, что $Q = 0$, a и κ постоянны и рассматривая только одно пространственное переменное (как в случае потока тепла вдоль тонкого стержня), получаем

$$\begin{aligned} \frac{\partial u}{\partial t} - \sigma \frac{\partial^2 u}{\partial x^2} &= 0, \quad \sigma = \text{const} > 0, \\ u &= u(x, t), \\ u(x, 0) &= u_0(x) \quad (\text{заданная функция}). \end{aligned} \quad (8.3)$$

Предположим, как и в гл. 1, что граничные условия заменены условиями периодичности.

Большинство обычных четырех- и шеститочечных разностных уравнений для уравнения (8.3) содержится в схеме

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} - \sigma \frac{\theta (\delta^2 u)_j^{n+1} + (1 - \theta) (\delta^2 u)_j^n}{(\Delta x)^2} = 0, \quad (8.4)$$

где θ является неотрицательной постоянной; значение $\theta = 0$ дает четырехточечную явную схему с взятой вперед разностью по времени. Другие значения θ дают неявные схемы; $\theta = 1/2$ дает шеститочечную схему с центральной разностью по времени, а $\theta = 1$ — четырехточечную схему с взятой назад разностью по времени. Как и в предыдущих главах, символ δ означает центральную разность по пространственному переменному. Простой алгоритм для решения таких неявных уравнений приводится ниже (в § 8.5).

Условие устойчивости для системы (8.4) было выведено в гл. 1 (см. неравенство (1.24)), но оно может быть получено также из общей теории, изложенной в гл. 4. Матрица перехода $G(\Delta t, k)$ содержит в данном случае только один элемент (потому что имеется только одно зависимое переменное) и является в действительности множителем перехода, обозначенным в гл. 1 через $\xi(m)$ и определенным формулой (1.23). Необходимое и достаточное условие для устойчивости, данное в гл. 4, совпадает и согласуется с условием, найденным в гл. 1, а именно ¹⁾

$$\begin{aligned} 2\sigma \Delta t / (\Delta x^2) &\leq 1 / (1 - 2\theta) \quad \text{при} \quad 0 < \theta \leq 1/2, \\ \text{ограничений нет} &\quad \text{при} \quad 1/2 \leq \theta \leq 1. \end{aligned}$$

В соответствии с результатами гл. 1 погрешность аппроксимации определяется формулой

$$e[u] = \frac{u_j^{n+1} - u_j^n}{\Delta t} - \sigma \frac{\theta (\delta^2 u)_j^{n+1} + (1 - \theta) (\delta^2 u)_j^n}{(\Delta x)^2},$$

где $u = u(x, t)$ есть точное решение дифференциального уравнения. Разлагая входящие в эту формулу функции в ряды Тейлора в окрестности точки $(x = j\Delta x, t = n\Delta t)$, находим, что $e[u] = O(\Delta t) + O[(\Delta x)^2]$. Так как эта величина стремится к нулю при неограниченном измельчении сетки, то из рассуждений гл. 4 и 7 следует, что условие согласованности разностного уравнения (8.4) с уравнением теплопроводности выполнено.

¹⁾ Условная устойчивость имеет место также при $\theta < 0$, а безусловная устойчивость — при $\theta > 1$, но, вообще говоря, эти значения θ не представляют интереса.

Если функция $u(x, t)$ — точное решение уравнения теплопроводности — является достаточно гладкой, то, учитывая, что $\partial^2 u / \partial t^2 = \sigma^2 \partial^4 u / \partial x^4$, погрешность аппроксимации можно записать в виде

$$e[u] = \sigma \frac{\partial^4 u}{\partial x^4} [\sigma \Delta t (1/2 - \theta) - 1/12 (\Delta x)^2] + O[(\Delta t)^2] + O[(\Delta x)^4]. \quad (8.5)$$

Поэтому, если постоянные θ , Δt и Δx выбраны так, что величина, стоящая в квадратных скобках в правой части равенства (8.5), равна нулю, то аппроксимация достигает более высокого порядка точности, чем в других случаях. При $0 \leq \theta < 1/2$ это имеет место, когда Δt взято равным одной трети максимального значения, допускаемого условием устойчивости. Очевидно также, что высшая степень точности получается, вообще говоря, не при центрировании разности по времени, т. е. при $\theta = 1/2$, а при $\theta = 1/2 - (\Delta x)^2 / 12 \sigma \Delta t$ (для этого значения θ аппроксимация устойчива). Саульев [1958] показал, что в частном случае, когда $(\Delta x)^2 / \sigma \Delta t = \sqrt{20}$, погрешность аппроксимации уменьшается до $O[(\Delta x)^6]$.

Для задач с переменными коэффициентами этот метод получения повышенной точности может быть несколько обобщен. Если $\sigma = \sigma(x, t)$ изменяется при изменении x и t , то оценка (8.5) для погрешности аппроксимации уже не является строго корректной, но если функция $\sigma(x, t)$ меняется медленно, то заметное повышение точности может быть по-прежнему получено при

$$\theta = \theta(x, t) = \frac{1}{2} - \frac{(\Delta x)^2}{12 \sigma(x, t) \Delta t}.$$

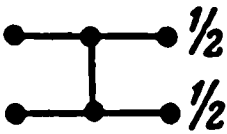
В этом случае более удобно переписать разностное уравнение в эквивалентной форме (включенной в табл. 8.1 под номером 12), из которой θ исключено, а далее преобразовать ее, как указано в § 3 гл. 8.

Таблица 8.1 содержит различные разностные схемы для простого уравнения диффузии. На расположенных слева мнемонических диаграммах, указывающих способ образования разностей, отмечены точки сетки на плоскости x, t , используемые при применении формулы; ось t направлена вертикально вверх, а ось x — горизонтально вправо. Если три узла связаны горизонтальной прямой, то они используются для образования второй пространственной разности; если два узла связаны вертикальной прямой, то они используются для образования разности по времени. Если имеются два или несколько множеств точек того и другого рода, то используемые веса указаны справа или снизу.

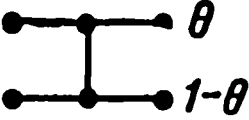
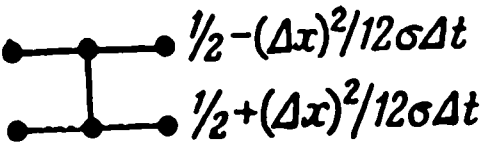
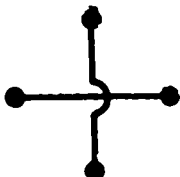
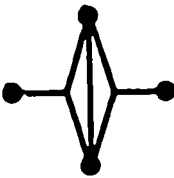
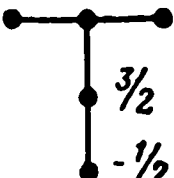
Таблица 8.I

Конечноразностные аппроксимации для уравнения

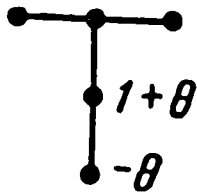
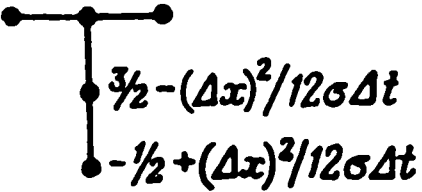
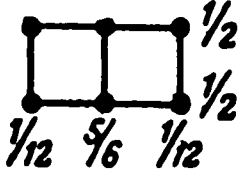
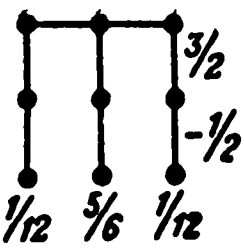
$$\frac{\partial u}{\partial t} = \sigma \frac{\partial^2 u}{\partial x^2}, \quad \sigma = \text{const} > 0$$

Расположение узлов разностной схемы	Разностные уравнения и погрешность аппроксимации
	1. $\frac{u_j^{n+1} - u_j^n}{\Delta t} = \sigma \frac{(\delta^2 u)_j^n}{(\Delta x)^2},$ $e = O(\Delta t) + O[(\Delta x)^2],$ явная схема, устойчива, если $\sigma \Delta t / (\Delta x)^2 = \text{const} \leq 1/2$ при $\Delta t, \Delta x \rightarrow 0$.
	2. $\frac{u_j^{n+1} - u_j^n}{\Delta t} = \sigma \frac{(\sigma^2 u)_j^n + (\delta^2 u)_j^{n+1}}{2(\Delta x)^2}$ (Кранк и Николсон [1947]), $e = O[(\Delta t)^2] + O[(\Delta x)^2],$ неявная схема, всегда устойчива.
	3. $\frac{u_j^{n+1} - u_j^n}{\Delta t} = \sigma \frac{(\sigma^2 u)_j^{n+1}}{(\Delta x)^2}$ (Лаасонен [1940]), $e = O(\Delta t) + O[(\Delta x)^2],$ неявная схема, всегда устойчива.
 (Частный случай)	4. Разностное уравнение такое же, как в случае 1, но $\sigma \Delta t / (\Delta x)^2 = 1/6,$ $e = O[(\Delta t)^2] = O[(\Delta x)^4],$ частный вариант случая 1; схема устойчива.

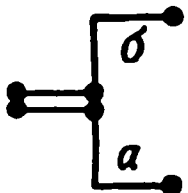
Продолжение

Расположение узлов разностной схемы	Разностные уравнения и погрешность аппроксимации
	5. $\frac{u_j^{n+1} - u_j^n}{\Delta t} = \sigma \frac{\theta (\sigma^2 u)_j^{n+1} + (1 - \theta) (\sigma^2 u)_j^n}{(\Delta x)^2},$ где $\theta = \text{const}$, $0 \leq \theta \leq 1$, $e = O(\Delta t) + O[(\Delta x)^2]$, при $0 \leq \theta < 1/2$ устойчива, если $\sigma \Delta t / (\Delta x)^2 = \text{const} \leq 1/(2 - 4\theta)$, при $1/2 \leq \theta \leq 1$ всегда устойчива, включает схемы 1 — 4 как частные случаи.
	6. Разностное уравнение такое же, как в случае 5, но $\theta = 1/2 - (\Delta x)^2 / 12 \sigma \Delta t$, $e = O[(\Delta t)^2] = O[(\Delta x)^4]$, схема устойчива.
	7. $\frac{u_j^{n+1} - u_j^{n-1}}{2 \Delta t} = \sigma \frac{(\delta^2 u)_j^n}{(\Delta x)^2},$ схема всегда неустойчива.
	8. $\frac{u_j^{n+1} - u_j^{n-1}}{2 \Delta t} = \sigma \frac{u_{j+1}^n - u_j^{n+1} - u_j^{n-1} + u_{j-1}^n}{(\Delta x)^2},$ где $\Delta t / \Delta x \rightarrow 0$ при $\Delta t, \Delta x \rightarrow 0$, $e = O(\Delta t) + O[(\Delta x)^2] + O(\Delta t / \Delta x)^2$ (Дюфорт и Франкел [1953]), явная схема, всегда устойчива.
	9. $\frac{3}{2} \frac{u_j^{n+1} - u_j^n}{\Delta t} - \frac{1}{2} \frac{u_j^n - u_j^{n-1}}{\Delta t} = \sigma \frac{(\delta^2 u)_j^{n+1}}{(\Delta x)^2},$ $e = O[(\Delta t)^2] + O[(\Delta x)^2],$ схема всегда устойчива.

Продолжение

Расположение узлов разностной схемы	Разностные уравнения и погрешность аппроксимации
	10. $(1 + \theta) \frac{u_j^{n+1} - u_j^n}{\Delta t} - \theta \frac{u_j^n - u_j^{n-1}}{\Delta t} = \sigma \frac{(\delta^2 u)_j^{n+1}}{(\Delta x)^2},$ где $\theta = \text{const} \geq 0$, $\sigma \Delta t / (\Delta x)^2 = \text{const}$, $e = O(\Delta t) + O[(\Delta x)^2]$, схема всегда устойчива, включает схемы 3 и 9 как частные случаи.
	11. Разностное уравнение такое же, как в случае 10, но при $\theta = 1/2 + (\Delta x)^2 / (12\sigma \Delta t)$, $e = O[(\Delta t)^2] = O[(\Delta x)^4]$, схема всегда устойчива.
	12. $\frac{1}{12} \frac{u_{j+1}^{n+1} - u_{j+1}^n}{\Delta t} + \frac{5}{6} \frac{u_j^{n+1} - u_j^n}{\Delta t} +$ $+ \frac{1}{12} \frac{u_{j-1}^{n+1} - u_{j-1}^n}{\Delta t} = \sigma \frac{(\delta^2 u)_j^{n+1} + (\delta^2 u)_j^n}{2(\Delta x)^2},$ $e = O[(\Delta t)^2] + O[(\Delta x)^4],$ схема всегда устойчива, совпадает со схемой 6.
	13. $\frac{1}{12} \frac{\frac{3}{2} u_{j+1}^{n+1} - 2u_{j+1}^n + \frac{1}{2} u_{j+1}^{n-1}}{\Delta t} +$ $+ \frac{5}{6} \frac{\frac{3}{2} u_j^{n+1} - 2u_j^n + \frac{1}{2} u_j^{n-1}}{\Delta t} +$ $+ \frac{1}{12} \frac{\frac{3}{2} u_{j-1}^{n+1} - 2u_{j-1}^n + \frac{1}{2} u_{j-1}^{n-1}}{\Delta t} =$ $= \frac{\sigma (\delta^2 u)_j^{n+1}}{(\Delta x)^2},$ $e = O[(\Delta t)^2] + O[(\Delta x)^4]$ схема всегда устойчива.

Продолжение

Расположение узлов разностной схемы	Разностные уравнения и погрешность аппроксимации
	14а. $u_j^{n+1} - u_j^n = \alpha (u_{j+1}^n - u_j^n - u_j^{n+1} + u_{j-1}^{n+1}),$ 14б. $u_j^{n+2} - u_j^{n+1} = \alpha (u_{j+1}^{n+2} - u_j^{n+2} - u_j^{n+1} + u_{j-1}^{n+1}),$ где $\alpha = \sigma \Delta t / (\Delta x)^2$ (Саульев [1957] — см. текст).

Возможны некоторые сравнения. Схемы 2 и 9 формально имеют одинаковый порядок точности; однако для гладких начальных функций предпочтительнее схема 2, потому что коэффициент при $(\Delta t)^2$ в ее погрешности аппроксимации меньше, тогда как для быстро меняющихся или разрывных начальных данных предпочтительнее схема 9, так как она быстрее гасит коротковолновые компоненты, для которых $k\Delta x \approx \pi$. Аналогичное соотношение имеет место между схемами 6 и 11 и между схемами 12 и 13.

В методе Саульева [1957] для двух последовательных шагов по времени применяются различные уравнения, как это показано в схеме 14 табл. 8. I. По существу эта схема является явной: каждое уравнение системы

$$u_j^{n+1} - u_j^n = \alpha (u_{j+1}^n - u_j^n - u_j^{n+1} + u_{j-1}^{n+1}) \quad (j = 1, \dots, J-1)$$

содержит две из неизвестных величин $u_0^{n+1}, \dots, u_{J-1}^{n+1}$, но u_0^{n+1} находится из левого граничного условия, так что уравнения могут быть последовательно решены для $j = 1, \dots, J-1$. Аналогично из второй системы уравнений можно последовательно получить u_j^{n+2} для $j = J-1, J-2, \dots, 1$. Для такой двухшаговой по времени схемы легко находится множитель перехода, равный

$$g = \frac{[1 - \alpha(1 - \cos \beta)]^2 + \alpha^2 \sin^2 \beta}{[1 + \alpha(1 - \cos \beta)]^2 + \alpha^2 \sin^2 \beta},$$

где $\alpha = \sigma \Delta t / (\Delta x)^2$, $\beta = k\Delta x$; так как $1 - \cos \beta \geq 0$, множитель g принадлежит отрезку $[-1, 1]$, и поэтому схема безусловно устойчива. Оценку порядка точности можно получить путем исключения u_j^{n+1} ; это дает вполне неявное уравнение

$$u^{n+2} - (\alpha + \alpha^2) \delta^2 u^{n+2} = u^n + (\alpha - \alpha^2) \delta^2 u^n.$$

Если пренебречь величиной α^2 , то получится схема Кранка и Николсона, погрешность аппроксимации которой равна $O[(\Delta t)^2] + O[(\Delta x)^2]$; вклад членов, содержащих α^2 , в $(u^{n+2} - u^n)/(2\Delta t)$ равен

$$O\left([\alpha \Delta x]^2 \frac{\partial^3 u}{\partial x^2 \partial t}\right) = O\left(\frac{(\Delta t)^2}{(\Delta x)^2} \frac{\partial^3 u}{\partial x^2 \partial t}\right).$$

При любом фиксированном α погрешность аппроксимации равна $O(\Delta t)$; если же Δt и Δx изменяются независимо, то схема Саульева, подобно схеме Дюфорта и Франкела, будет согласована с уравнением теплопроводности лишь тогда, когда $\Delta t/\Delta x \rightarrow 0$ при измельчении сетки.

§ 8.3. Переменные коэффициенты

Большинство схем, указанных в табл. 8.1, может быть обобщено на случай переменных коэффициентов¹⁾. Если обобщаются схемы, обладающие повышенной точностью, то следует заботиться о том, чтобы не произошло потери точности, за исключением тех случаев, когда коэффициенты близки к постоянным. Для иллюстрации покажем, как обобщается схема 12. Сначала рассмотрим уравнение

$$\frac{\partial u}{\partial t} = \sigma \frac{\partial^2 u}{\partial x^2}, \quad \sigma = \sigma(x) \geq a > 0, \quad (8.6)$$

и допустим, что $\sigma(x)$ является дважды непрерывно дифференцируемой функцией. Удобно сначала ввести пространственные разностные отношения

$$\frac{\delta^2 u}{(\Delta x)^2} = \frac{\partial^2 u}{\partial x^2} + \frac{(\Delta x)^2}{12} \frac{\partial^4 u}{\partial x^4} + O[(\Delta x)^4],$$

откуда

$$\begin{aligned} \frac{(\delta^2 u)_j}{(\Delta x)^2} &= \left(\frac{1}{\sigma} \frac{\partial u}{\partial t}\right)_j + \frac{(\Delta x)^2}{12} \left[\frac{\partial^2}{\partial x^2} \left(\frac{1}{\sigma} \frac{\partial u}{\partial t}\right)\right]_j + O[(\Delta x)^4] = \\ &= \left(\frac{1}{\sigma} \frac{\partial u}{\partial t}\right)_j + \frac{(\Delta x)^2}{12} \left[\left\{\frac{\delta^2 \left(\frac{1}{\sigma} \frac{\partial u}{\partial t}\right)}{(\Delta x)^2}\right\}_j + O[(\Delta x)^2]\right] + O[(\Delta x)^4] = \\ &= \frac{5}{6} \left(\frac{1}{\sigma} \frac{\partial u}{\partial t}\right)_j + \frac{1}{12} \left(\frac{1}{\sigma} \frac{\partial u}{\partial t}\right)_{j+1} + \frac{1}{12} \left(\frac{1}{\sigma} \frac{\partial u}{\partial t}\right)_{j-1} + O[(\Delta x)^4]. \end{aligned}$$

¹⁾ Схема 4 не может быть обобщена таким образом, поскольку специальное условие $\sigma \Delta t/(\Delta x)^2 = 1/6$ требует, чтобы σ было постоянным.

Разностные отношения по времени теперь образуются так же, как и раньше, и мы получаем в результате формулу

$$\begin{aligned} \frac{1}{12} \frac{u_{j+1}^{n+1} - u_{j+1}^n}{\sigma_{j+1} \Delta t} + \frac{5}{6} \frac{u_j^{n+1} - u_j^n}{\sigma_j \Delta t} + \frac{1}{12} \frac{u_{j-1}^{n+1} - u_{j-1}^n}{\sigma_{j-1} \Delta t} = \\ = \frac{(\delta^2 u)_j^{n+1} + (\delta^2 u)_j^n}{2(\Delta x)^2} + O[(\Delta t)^2] + O[(\Delta x)^4]. \end{aligned} \quad (8.7)$$

Для уравнения более общего вида

$$\frac{\partial u}{\partial t} = \mu \frac{\partial}{\partial x} v \frac{\partial u}{\partial x}, \quad \text{где} \quad \begin{cases} \mu = \mu(x) \geq a_1 > 0, \\ v = v(x) \geq a_2 > 0, \end{cases}$$

формулы несколько усложняются, и мы их не будем выписывать. Заметим только, что такие формулы можно легко получить заменой переменных

$$y = \int \frac{1}{v(x)} dx,$$

после которой уравнение принимает вид

$$\frac{\partial u}{\partial t} = \frac{\mu}{v} \frac{\partial^2 u}{\partial y^2},$$

т. е. имеет такую же форму, как уравнение (8.6), и может решаться таким же образом. Для исходного переменного x эта схема имеет неравномерную сетку.

Формулы такого типа использовались в вычислениях с диффузионно-возрастным уравнением Ферми для ядерных реакторов и дали возможность получить точные результаты с меньшим объемом вычислений, чем требуется при более простых формулах.

Что касается устойчивости этих разностных схем с переменными коэффициентами, удовлетворяющими условию Липшица, то в § 5.3 было показано, что она имеет место для явных схем, множитель перехода которых удовлетворяет условию (5.22), а именно

$$|G(x, \xi)| \leq 1 - \delta |\xi|^2 \text{ при всех } x \text{ и } |\xi| \leq \pi,$$

где $\xi = k\Delta x$ и постоянная $\delta > 0$; там же вкратце говорилось о некоторых результатах для неявных схем. Чтобы понять, что означает это условие, рассмотрим схему 5 из табл. 8.1, для которой

$$G = \frac{1 - 4\alpha(1 - \theta) \sin^2(\xi/2)}{1 + 4\alpha\theta \sin^2(\xi/2)},$$

где $\alpha(x, \Delta t) = \sigma(x) \Delta t / (\Delta x)^2$. Следовательно,

$$-1 + \frac{2 - 4\alpha(1 - 2\theta) \sin^2(\xi/2)}{1 + 4\alpha \sin^2(\xi/2)} = G = 1 - \frac{4\alpha \sin^2(\xi/2)}{1 + 4\alpha\theta \sin^2(\xi/2)}$$

и указанное условие выполняется, если i) величина $\alpha(x, \Delta t)$ равномерно по x ограничена снизу положительной постоянной при $\Delta t \rightarrow 0$ и ii) $2\alpha(x, \Delta t)(1 - 2\theta) \leq \gamma < 1$. Так как γ может быть взято сколь угодно близким к единице, то это практически не накладывает ограничения на шаг Δt , с которым можно проводить вычисления (сравните со сказанным в следующем параграфе).

Здесь стоит также отметить, что доказательство устойчивости для схем с переменными коэффициентами, приведенное в § 5.3, использует явность схемы лишь в незначительной степени. Если в сумме $u^{n+1}(x) = \sum_{(\beta)} C^\beta u^n(x + \beta \Delta x)$ абсолютно сходятся как C^β , так и их постоянные Липшица, то доказательство не будет существенно отличаться от того, которое приведено в § 3 гл. 5.

§ 8.4. Влияние на устойчивость членов низшего порядка

Теперь мы рассмотрим уравнение

$$\frac{\partial u}{\partial t} = \sigma \frac{\partial^2 u}{\partial x^2} + a \frac{\partial u}{\partial x} + bu, \quad (8.8)$$

где σ , a и b постоянны, причем $\sigma > 0$. При построении конечно-разностного уравнения каждый член уравнения (8.8) может быть заменен разностными отношениями различным образом, и возможные комбинации этих способов весьма многочисленны. Мы дадим два примера. Эти (и другие) примеры подтверждают вывод о том, что по крайней мере для задач диффузии устойчивость практически не зависит от членов низшего порядка. Слово «практически» имеет тот смысл, что в некоторых случаях ограничение $\Delta t \leq \dots$ должно быть заменено ограничением $\Delta t < \dots$ без каких-либо других изменений. (В § 5.3 это было показано для самых общих явных схем, аппроксимирующих уравнение (8.8), с помощью тех же по существу рассуждений, которые применяются в приводимых ниже примерах.) Однако наличие членов низшего порядка может явиться причиной уменьшения Δt : например, большое значение b в уравнении (8.8) приводит к быстрому изменению решения u , даже если u не зависит от x , и тогда малость Δt необходима точно также, как и для конечноразностной аппроксимации обыкновенного дифференциального уравнения $du/dt = bu$.

В качестве первого примера рассмотрим пространственно центрированную схему

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \sigma \frac{\theta (\delta^2 u)_j^{n+1} + (1 - \theta) (\delta^2 u)_j^n}{(\Delta x)^2} + a \frac{u_{j+1}^n - u_{j-1}^n}{2 \Delta x} + bu_j^n, \quad (8.9)$$

где $\theta = \text{const}$, $0 \leq \theta \leq 1$. Пусть $\alpha = \sigma \Delta t / (\Delta x)^2$ остается постоянным при $\Delta t \rightarrow 0$ и $\Delta x \rightarrow 0$, так что $\Delta x = \sqrt{\sigma \Delta t / \alpha}$. Множитель перехода $G = G(\Delta t, k)$ равен

$$G = \frac{1 - 4\alpha(1 - \theta) \sin^2 \beta + ia \sqrt{\alpha \Delta t / \sigma} \sin 2\beta + b \Delta t}{1 + 4\alpha\theta \sin^2 \beta},$$

где $\beta = k\Delta x/2$. Поэтому

$$|G| = \frac{\sqrt{[1 - 4\alpha(1 - \theta) \sin^2 \beta + b \Delta t]^2 + a^2 (\alpha \Delta t / \sigma) \sin^2 2\beta}}{1 + 4\alpha\theta \sin^2 \beta},$$

что мы перепишем в следующем виде:

$$|G| = \sqrt{f_0(\beta) + \Delta t f_1(\beta) + (\Delta t)^2 f_2(\beta)},$$

где функции $f_0(\beta)$, $f_1(\beta)$ и $f_2(\beta)$ ограничены. Очевидно, нужно потребовать, чтобы $|f_0(\beta)| \leq 1$, и это приводит к тому же самому условию, что и в случае $a = b = 0$, а именно

$$\begin{aligned} \alpha &\leq 1/(2 - 4\theta) && \text{при } 0 \leq \theta < 1/2, \\ \text{ограничений на } \Delta t &\text{ нет при } 1/2 \leq \theta \leq 1. \end{aligned} \quad (8.10)$$

Если это условие выполняется и если m_1 и m_2 суть максимальные значения $|f_1(\beta)|$ и $|f_2(\beta)|$ при изменении β , то мы имеем

$$|G| \leq \sqrt{1 + m_1 \Delta t + m_2 (\Delta t)^2} = 1 + O(\Delta t),$$

и поэтому условие устойчивости оказывается таким же, как и ранее.

В качестве второго примера, быть может, менее удовлетворительного с точки зрения погрешности аппроксимации, рассмотрим (нецентрированную) схему с нецентрированной первой разностью по пространству

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \sigma \frac{\theta (\delta^2 u)_j^{n+1} + (1 - \theta) (\delta^2 u)_j^n}{(\Delta x)^2} + a \frac{\Delta_- u_j^n}{\Delta x} + b u_j^n, \quad (8.11)$$

где, как и в предыдущих главах, $\Delta_- u_j^n = u_j^n - u_{j-1}^n$, и снова предположим, что $\sigma \Delta t / (\Delta x)^2 = \alpha$ имеет фиксированное значение при $\Delta t \rightarrow 0$ и $\Delta x \rightarrow 0$. Сначала рассмотрим случай $\alpha \geq 0$.

Теперь мы имеем

$$G = \frac{1 - 4\alpha(1 - \theta) \sin^2 \beta + a \sqrt{\frac{\alpha \Delta t}{\sigma}} (2 \sin^2 \beta + i \sin 2\beta) + b \Delta t}{1 + 4\alpha\theta \sin^2 \beta}. \quad (8.12)$$

Снова ясно, что (8.10) является по меньшей мере необходимым условием устойчивости. Покажем, что оно является также и достаточным.

Устойчивость связана только с тем, что происходит в пределе при $\Delta t, \Delta x \rightarrow 0$. Таким образом, необходимо рассматривать лишь малые Δt . Допустим поэтому, что

$$2a \sqrt{\alpha \Delta t / \sigma} \leq 4a, \text{ т. е. } \Delta x \leq 2\sigma/a.$$

Если условие (8.10) выполнено, то мы имеем

$$\begin{aligned} -1 &\leq \frac{1 - 4\alpha(1 - \theta) \sin^2 \beta}{1 + 4\alpha\theta \sin^2 \beta} \leq \\ &\leq \frac{1 - 4\alpha(1 - \theta) \sin^2 \beta + 2a \sqrt{\alpha \Delta t / \sigma} \sin^2 \beta}{1 + 4\alpha\theta \sin^2 \beta} \leq 1. \end{aligned} \quad (8.13)$$

Обозначив третий член в этом неравенстве через A , получим

$$|G|^2 = \left(A + \frac{b \Delta t}{1 + 4\alpha\theta \sin^2 \beta} \right)^2 + \left(\frac{a \sqrt{\alpha \Delta t / \sigma} \sin 2\beta}{1 + 4\alpha\theta \sin^2 \beta} \right)^2,$$

и поэтому

$$|G|^2 \leq 1 + O(\Delta t) \text{ равномерно относительно } \beta,$$

что и доказывает устойчивость.

Рассмотрим теперь противоположный случай $a \leq 0$. Множитель перехода G по-прежнему дается соотношением (8.12), но так как $a \leq 0$, знак неравенства между вторым и третьим членами неравенства (8.13) нужно заменить противоположным, и поэтому существует опасение, что может оказаться $A < -1$. Во избежание этого нужно изменить первую строку условия (8.10) и потребовать, чтобы

$$\begin{aligned} \alpha &< 1/(2 - 4\theta) \quad \text{при} \quad 0 \leq \theta < 1/2, \\ \text{ограничений на } \alpha &\text{ нет при} \quad 1/2 \leq \theta \leq 1. \end{aligned} \quad (8.14)$$

(Это отличается от (8.10) только тем, что теперь не допускается равенства $\alpha = 1/(2 - 4\theta)$.) В каждом случае для заданных α и θ существует такая постоянная B , что

$$-1 < B \leq \frac{1 - 4\alpha(1 - \theta) \sin^2 \beta}{1 + 4\alpha\theta \sin^2 \beta}.$$

Так как A отличается от этой последней величины только на отрицательное слагаемое, пропорциональное $\sqrt{\Delta t}$, то ясно, что, выбирая Δt достаточно малым, можно добиться того, чтобы $-1 \leq A$. Тогда $-1 \leq A \leq 1$, а отсюда дальнейшие выводы делаются точно так же, как и в случае $a \geq 0$.

Если в схеме (8.11) мы возьмем разность вперед $\Delta_+ u_j^n$ вместо $\Delta_- u_j^n$, то результат будет тем же самым, за исключением того, что случаи $a \geq 0$ и $a \leq 0$ поменяются ролями.

§ 8.5. Решение неявных уравнений

Если уравнение

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x} \sigma \frac{\partial u}{\partial x}, \quad \sigma = \sigma(x) > 0,$$

(мы берем типичный случай) аппроксимируется конечноразностной схемой

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\theta [\delta(\sigma \delta u)]_j^{n+1} + (1 - \theta) [\delta(\sigma \delta u)]_j^n}{(\Delta x)^2},$$

аналогичной схеме 5 табл. 8. I, то уравнения, которые нужно решать на каждом шагу по времени, будут иметь вид

$$-\theta \alpha_{j+1/2} u_{j+1}^{n+1} + (1 + \theta \alpha_{j+1/2} + \theta \alpha_{j-1/2}) u_j^{n+1} - \theta \alpha_{j-1/2} u_{j-1}^{n+1} = \dots, \\ j = 1, 2, \dots,$$

где черточки в правой части равенства означают известные величины, а $\alpha = \alpha(x)$ стоит вместо $\sigma(x) \Delta t / (\Delta x)^2$. Эти уравнения должны быть решены при некоторых граничных условиях, соответствующих, например, $j = 0$ и $j = J$. Перепишем эти уравнения в виде

$$-A_j u_{j+1} + B_j u_j - C_j u_{j-1} = D_j, \quad (8.15)$$

где A_j, B_j, C_j — сокращенные обозначения коэффициентов, причем верхний индекс $n + 1$ опущен. Для уравнения $\partial u / \partial t = \sigma \partial^2 u / \partial x^2$ коэффициенты будут несколько иными. В качестве второго примера рассмотрим аппроксимацию этого уравнения разностным уравнением (8.7) (схема 12 табл. 8. I), которое эквивалентно тому случаю, когда θ в первом примере будет переменным и в каждой точке сетки будет принимать значение $\theta = 1/2 - (\Delta x)^2 / 12 \sigma \Delta t$, характеризующее схему 6 табл. 8. I; тогда

$$B_j = 1 + 5/(6\alpha_j), \quad A_j = 1/2 - 1/(12\alpha_{j+1}), \quad C_j = 1/2 - 1/(12\alpha_{j-1}).$$

Если предположить, что $\theta > 0$ в первом примере и $\alpha_j > 1/6$ для всех j во втором, то для обоих примеров

$$A_j > 0, \quad B_j > 0, \quad C_j > 0 \quad (8.16)$$

и, кроме того, неравенство

$$B_j > A_j + C_j \quad (8.17)$$

выполняется с некоторым запасом.

Чтобы проиллюстрировать метод решения рекуррентной системы вида (8.15), предположим, что граничные условия имеют вид

$$u_0 = 0, \quad (8.18)$$

$$u_J = 0. \quad (8.19)$$

Читатель заметит, что при небольших изменениях метод может быть приспособлен и к другим граничным условиям.

Уравнения, подлежащие решению, линейны, и можно было бы применить один из стандартных методов решения линейных систем. Но эта система уравнений имеет весьма специальный вид, так как все элементы соответствующей матрицы (за исключением элементов, расположенных на трех диагоналях) обращаются в нуль, а для этих последних выполняются неравенства (8.16) и (8.17). Метод, который мы опишем для таких систем, является весьма эффективным и очень удобным для машинных вычислений¹⁾.

Заметим сначала, что *не является* хорошим следующий (иногда применяемый) метод. Берут некоторое решение $u^{(1)}$ уравнений (8.15), удовлетворяющее левому граничному условию (8.18), и некоторое решение $u^{(2)}$ соответствующих однородных уравнений (т. е. системы (8.15) при $D_j = 0$), также удовлетворяющее левому граничному условию, а затем составляют их линейную комбинацию, так, чтобы удовлетворить и правому граничному условию. Причина, по которой этот метод плох²⁾, состоит в том, что в общем случае $u^{(1)}$ и $u^{(2)}$ растут очень быстро, приблизительно по экспонентам, когда j возрастает от 0 до J , и поэтому истинное решение должно получаться как разность двух очень больших и почти одинаковых величин. Это требует не только использования плавающей запятой, но и многократного повышения точности счета. В методе, описываемом ниже, все величины остаются в разумных границах.

При описании метода мы тем не менее рассмотрим множество S решений разностного уравнения (8.15), удовлетворяющих левому граничному условию (8.18): S является однопараметрическим семейством решений, потому что каждым значением u_j при $j = 1$ такое решение определяется однозначно. Действительно, найдем два множества величин E_j и F_j , таких, что для каждого элемента множества S

$$u_j = E_j u_{j+1} + F_j. \quad (8.20)$$

¹⁾ Насколько нам известно, этот метод был применен независимо многими авторами; конечно, он является лишь специальной формой процесса исключения Гаусса.

²⁾ Подробно этот вопрос рассмотрен, например, в книге Годунова и Рябенского [1962], стр. 164—171. — Прим. перев.

В частности, если это должно выполняться для каждого члена данного семейства, то граничное условие (8.18) показывает, что должно быть

$$E_0 = 0, \quad F_0 = 0. \quad (8.21)$$

Позже мы увидим, что $E_j > 0$ для $j > 0$. Поэтому система (8.20) также имеет однопараметрическое семейство решений, так как u_1 может быть выбрано произвольно, а тогда система (8.20) однозначно определяет все u_j при $j > 1$. Наша задача состоит в таком определении E_j и F_j , чтобы это семейство совпало с семейством S .

Чтобы этого достигнуть, подставим $E_{j-1}u_j + F_{j-1}$ вместо u_{j-1} в уравнение (8.15). В результате получим соотношение между u_j и u_{j+1} , которое можно записать так:

$$u_j = \frac{A_j}{B_j - C_j E_{j-1}} u_{j+1} + \frac{D_j + C_j F_{j-1}}{B_j - C_j E_{j-1}}.$$

Если мы приравняем правую часть этого равенства правой части равенства (8.20) и вспомним, что этот результат должен быть справедлив для однопараметрического семейства величин u_{j+1} , то станет очевидным, что для этого должны выполняться соотношения

$$E_j = \frac{A_j}{B_j - C_j E_{j-1}}, \quad j \geq 1. \quad (8.22)$$

$$F_j = \frac{D_j + C_j F_{j-1}}{B_j - C_j E_{j-1}}, \quad j \geq 1. \quad (8.23)$$

С помощью равенств (8.21) — (8.23) можно вычислить E_j и F_j по индукции в направлении возрастания j ($j = 0, 1, 2, \dots, J-1$). Так как u_{j+1} для $j = J-1$ дается с помощью правого граничного условия (8.19), мы можем теперь с помощью (8.20) вычислить u_j по индукции в направлении убывания j ($j = J-1, J-2, \dots, 1$). Этим вычисление завершается.

Чтобы показать, что E_j и F_j хорошо расположены в шкале чисел, заметим сначала, что если $E_{j-1} \leq 1$, то в силу равенства (8.22) и неравенства (8.17)

$$0 < E_j \leq \frac{A_j}{B_j - C_j} < \frac{A_j}{A_j} = 1.$$

Так как $E_0 = 0$, то это означает, что E_j лежит между 0 и 1 для всех j . Кроме того, как заметил Муссман, если E_j и искомое решение u_j должным образом ограничены, то из равенств (8.20) очевидным образом следует, что и F_j соответственно ограничены.

Чтобы показать эффективность этого метода, заметим, что, помимо вычисления коэффициентов, требуется только три умно-

жения и два деления для каждой точки пространства на каждом шагу по времени. Это может быть противопоставлено процессу, рассмотренному для задач, в которых коэффициенты не зависят от времени, а именно обращению раз и навсегда матрицы системы уравнений (8.15), (8.18) и (8.19) и последующему применению этой матрицы для получения решения на каждом шагу по времени. Кроме этой работы по обращению матрицы требуется еще умножение матрицы на вектор, т. е. J умножений, для каждой точки пространства на каждом шагу по времени.

§ 8.6. Нелинейные задачи

Применение вышеизложенных идей к нелинейным задачам проиллюстрируем исследованием уравнения

$$\frac{\partial u}{\partial t} = \frac{\partial^2}{\partial x^2} (u^5), \quad (8.24)$$

для которого были выполнены некоторые пробные расчеты на машине УНИВАК.

Решение выбранной нами смешанной краевой задачи известно в аналитической форме, так что точность решения, полученного численным методом, можно оценить сравнением. Решение уравнения (8.24) в виде распространяющейся волны может быть получено как решение, в котором u зависит от x и t только в комбинации $x - vt$, где v постоянно. Общее решение этого типа неявно дается уравнением

$$\begin{aligned} & \frac{5}{4}(u - u_0)^4 + \frac{20}{3}u_0(u - u_0)^3 + 15u_0^2(u - u_0)^2 + \\ & + 20u_0^3(u - u_0) + 5u_0^4 \ln(u - u_0) = v(vt - x + x_0), \end{aligned} \quad (8.25)$$

где u_0 , x_0 и v постоянны. Это — волна, распространяющаяся вправо при $v > 0$; ее форма показана на рис. 8.1. Слева решение u близко к кривой четвертого порядка, а справа приближается к u_0 при возрастании x примерно по экспоненте, потому что при $u - u_0 \ll u_0$ логарифм является преобладающим членом в левой части равенства (8.25). Без потери общности мы можем взять $u_0 = 1$. Обозначим через $\psi(\xi)$ обратную функцию для левой части уравнения (8.25); тогда решением в виде распространяющейся волны является

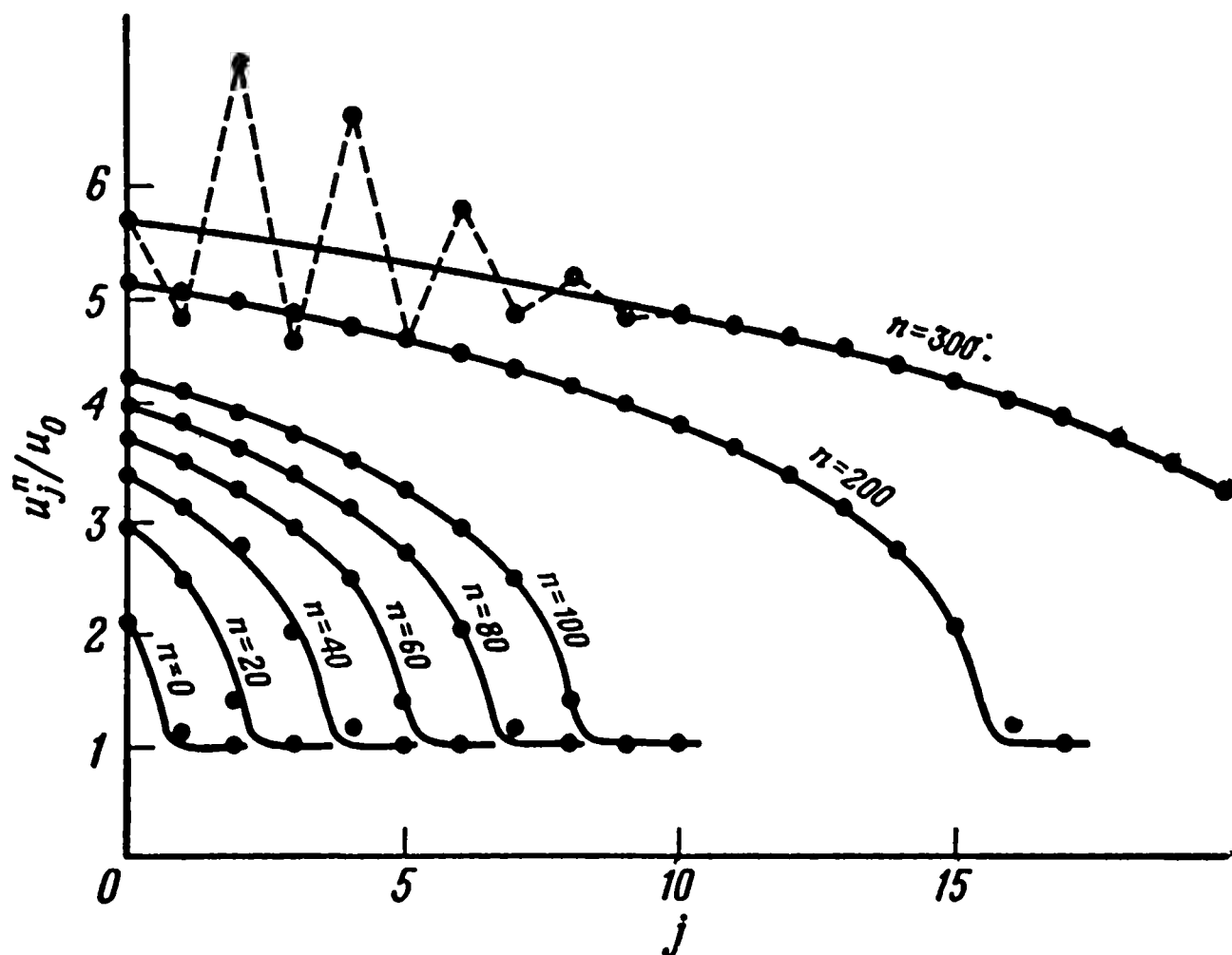
$$u = \psi(v(vt - x + x_0)),$$

а функция $\psi(\xi)$ может быть получена графически или с помощью метода Ньютона — Рафсона из уравнения (8.25). На-

начальные и граничные значения возьмем из этого решения, т. е.

$$\begin{aligned} \text{начальное условие} \quad & u(x, 0) = \psi(v(x_0 - x)), \\ \text{граничные условия:} \quad & \begin{cases} u(0, t) = \psi(v(vt + x_0)) \\ u(L, t) = \psi(v(vt - L + x_0)). \end{cases} \end{aligned} \quad (8.26)$$

Не нарушая общности, можно взять длину интервала $L = 1$. В силу единственности решения уравнения (8.24) численный метод, если он имеет некоторую точность, должен воспроизводить



Р и с. 8.1. Решение нелинейного уравнения $\partial u / \partial t = \partial^2 (u^5) / \partial x^2$, имеющее вид распространяющейся волны.

Сплошной линией изображено точное решение, найденное из уравнения (8.25), а штриховой — решение разностного уравнения (8.27) при $\theta = 0.4$ и при Δt и Δx , выбранных так, что $v \Delta t / \Delta x = 0.075$ и $5u_0^4 \Delta t / (\Delta x)^2 = 0.005$. Числа n показывают номера шагов по времени.

решение в виде распространяющейся волны при $0 \leq x \leq 1$, $t \geq 0$.

В качестве конечноразностной системы возьмем обобщение общей неявной системы 5 табл. 8. I, так чтобы можно было изучать точность и устойчивость для различных значений θ . Непосредственное обобщение разностного уравнения

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\theta [\delta^2 (u^5)]_j^{n+1} + (1 - \theta) [\delta^2 (u^5)]_j^n}{(\Delta x)^2}$$

было бы неуместным, потому что для него система уравнений, подлежащая решению, была бы нелинейной. Линеаризованный вариант уравнения может быть получен различными способами:

удобным для рассматриваемой задачи является вариант, получающийся аппроксимацией

$$(u^5)_j^{n+1} \approx (u^5)_j^n + 5 (u^4)_j^n (u_j^{n+1} - u_j^n).$$

Положив $w_j = u_j^{n+1} - u_j^n$, мы получим уравнения

$$\begin{aligned} \frac{w_j}{\Delta t} - \frac{5\theta}{(\Delta x)^2} \left[(u^4)_{j+1}^n w_{j+1} - 2 (u^4)_j^n w_j + (u^4)_{j-1}^n w_{j-1} \right] = \\ = \frac{(u^5)_{j+1}^n - 2 (u^5)_j^n + (u^5)_{j-1}^n}{(\Delta x)^2} \end{aligned} \quad (8.27)$$

вместе с граничными условиями

$$w_0 = u_0^{n+1} - u_0^n, \quad w_J = u_J^{n+1} - u_J^n,$$

в которых члены, стоящие в правой части, известны. Эти уравнения могут быть решены методом, изложенным в предыдущем параграфе; применение после этого уравнения $u_j^{n+1} = u_j^n + w_j$ завершает вычисления для данного шага по времени.

Задача была запрограммирована так, что начальные и граничные значения получались путем решения уравнения (8.25) методом Ньютона — Рафсона, а решение конечноразностных уравнений получалось, как указано выше. После каждых двадцати шагов по времени машина печатала столбцом полученные таким путем для $j = 0, 1, \dots, J$ значения u_j^n , а рядом — параллельным столбцом — точные значения, полученные путем решения уравнения (8.25) методом Ньютона — Рафсона. Образец такой записи воспроизводится в виде табл. 8. II.

Из этой таблицы видно, что совпадение, вообще говоря, оказывается хорошим. На рис. 8.1 графически показано несколько множеств таких значений для одной из подобных задач; они получены при $\Delta t / (\Delta x)^2 = 0,001$, $\theta = 0,4$. Решение в виде распространяющейся волны получается весьма точно, за исключением двух участков, а именно:

- 1) имеются слабые расхождения вблизи основания каждой волны, где точное решение меняется наиболее быстро;
- 2) на верхнем левом участке последние две кривые имеют колебания, свидетельствующие о появлении неустойчивости.

Скорость движения волны вправо, оцененная с помощью графика в точках с абсциссой, соответствующей примерно половине пути, совпадает с точным значением с точностью до долей процента. В общем, принимая во внимание грубость сетки, результат можно считать совершенно удовлетворительным.

Эвристически к вопросу об устойчивости можно подойти следующим образом: эффективным коэффициентом диффузии σ

Таблица 8.11

Часть численного решения

<i>l</i>	Теория	Эксперимент	Критерий устойчивости
00	04,9885291	04,9885812	00,0000000
01	04,8892890	04,8892689	01,1428989
02	04,7840560	04,7841171	01,0477001
03	04,6722668	04,6722014	00,9530582
04	04,5525326	04,5524622	00,8590384
05	04,4235425	04,4234839	00,7657584
06	04,2835565	04,2833949	00,6732604
07	04,1300680	04,1297196	00,5817199
08	03,9594097	03,9589363	00,4912986
09	03,7667231	03,7658970	00,4022574
10	03,5437619	03,5423956	00,3149394
11	03,2767332	03,2745918	00,2299584
12	02,9386536	02,9355146	00,1485191
13	02,4614235	02,4580572	00,0730170
14	01,5669351	01,4850563	00,0097235
15	01,0000000	01,0122755	00,0020943
16	01,0000000	01,0000249	00,0020000
17	01,0000000	01,0000000	00,0020000
18	01,0000000	01,0000000	00,0020000
19	01,0000000	01,0000000	00,0020000
20	01,0000000	01,0000000	00,0020000

для этой задачи является $5u^4$, что можно усмотреть, переписав дифференциальное уравнение в виде

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x} 5u^4 \frac{\partial u}{\partial x},$$

и поэтому σ зависит от x и t . Замечая, что неустойчивость, когда она возникает, проявляется в быстрых колебаниях локального характера, можно ожидать, что метод будет устойчивым до момента $t = t_0$ тогда и только тогда, когда

$$\frac{5u^4 \Delta t}{(\Delta x)^2} < \frac{1}{2 - 4\theta} \quad \text{при} \quad 0 \leq \theta < 1/2$$

$$\text{и при всех} \quad \begin{cases} 0 \leq x \leq L, \\ 0 \leq t \leq t_0, \end{cases} \quad (8.28)$$

ограничений нет при $1/2 \leq \theta \leq 1$.

Заметим, что для нелинейных задач устойчивость зависит не только от структуры конечноразностной системы, но также, вообще говоря, и от получаемого решения; при данном решении система может быть устойчивой для некоторых значений t и не быть таковой для других. Это утверждение можно сравнить с результатом Ф. Джона [1952] для явных схем, описанным в § 5.3.

В данной задаче при $\theta = 0,4$ и $\Delta t/(\Delta x)^2 = 0,001$ мы ожидаем, что устойчивость будет иметь место до тех пор, пока u (которое постоянно возрастает всюду в случае решения в виде распространяющейся волны) не достигнет значения $u = (500)^{1/4} \approx 4,7$, и что вскоре после этого будет развиваться неустойчивость. Именно это имело место при пробных вычислениях, что можно увидеть на рис. 8.1. Отметим, что для этой задачи при выбранной сетке явная система приводила бы к неустойчивости уже на 40-м шагу по времени вместо 250-го. Позднее были сделаны пробные расчеты с разными θ и Δt . Во всех случаях предсказания, полученные с помощью эвристического подхода к устойчивости подтвердились.

При практическом решении нелинейных дифференциальных уравнений необходимо, или по крайней мере в высшей степени желательно, чтобы машина постоянно следила бы за устойчивостью (если разностные уравнения не являются безусловно устойчивыми), проверяя, например, выполнение неравенства для Δt в условии (8.28) и либо останавливала счет, когда условие устойчивости не выполняется, либо автоматически меняла шаг Δt , чтобы восстановить устойчивость.

§ 8.7. Задачи с несколькими пространственными переменными

В качестве первого примера рассмотрим параболическое уравнение для функции $u = u(x, y, t)$

$$\frac{\partial u}{\partial t} = A \frac{\partial^2 u}{\partial x^2} + 2B \frac{\partial^2 u}{\partial x \partial y} + C \frac{\partial^2 u}{\partial y^2}, \quad (8.29)$$

где

$$A > 0, \quad C > 0, \quad AC - B^2 > 0. \quad (8.30)$$

Неравенства (8.30) обеспечивают параболичность уравнения; если они выполнены, то краевая задача для уравнения (8.29) поставлена корректно.

Для записи разностных уравнений введем сокращенные обозначения

$$u_{jl}^n = u(j \Delta x, l \Delta y, n \Delta t), \quad (8.31)$$

$$\begin{aligned} \Phi_{jl}^n = & \frac{A}{(\Delta x)^2} (u_{j+1,l}^n - 2u_{jl}^n + u_{j-1,l}^n) + \\ & + \frac{B}{2 \Delta x \Delta y} (u_{j+1,l+1}^n - u_{j-1,l+1}^n - u_{j+1,l-1}^n + u_{j-1,l-1}^n) + \\ & + \frac{C}{(\Delta y)^2} (u_{j,l+1}^n - 2u_{j,l}^n + u_{j,l-1}^n). \end{aligned} \quad (8.32)$$

Взяв постоянную θ ($0 \leq \theta \leq 1$), напомним разностную аппроксимацию для уравнения (8.29) в виде

$$\frac{u_{j,l}^{n+1} - u_{jl}^n}{\Delta t} = \theta \Phi_{jl}^{n+1} + (1 - \theta) \Phi_{jl}^n. \quad (8.33)$$

Если $\theta > 0$, то эти уравнения неявны; для каждого n нужно решить систему уравнений, аналогичную системе, получающейся при численном решении эллиптической краевой задачи для функции $u(x, y)$. Не существует простого обобщения алгоритма § 8.5 для решения этих уравнений, и, вообще говоря, приходится обращаться к методу релаксации. Раз это так, то может возникнуть мысль, что выгоднее было бы применять явные уравнения ($\theta = 0$) и брать Δt достаточно малым, чтобы эти уравнения были устойчивы. Однако разработаны очень эффективные релаксационные методы¹⁾, и u_{jl}^n обычно обеспечивает превосходное первое приближение для u_{jl}^{n+1} . При численных расчетах ядерных реакторов²⁾ было найдено, что неявные схемы целесообразно применять для решения двумерного диффузионно-возрастного уравнения, которое сходно по форме с уравнением (8.29).

Подстановка $\exp[ik_1 x + ik_2 y]$ вместо $u^n(x, y)$ в выражение Φ_{jl}^n , определяемое формулой (8.32), после умножения на Δt дает

$$\Delta t \Phi_{jl}^n = \psi \cdot \exp[ik_1 j \Delta x + ik_2 l \Delta y],$$

где

$$\begin{aligned} \psi = & -2 \left[\frac{A \Delta t}{(\Delta x)^2} (1 - \cos k_1 \Delta x) + \right. \\ & \left. + \frac{B \Delta t}{\Delta x \Delta y} \sin k_1 \Delta x \sin k_2 \Delta y + \frac{C \Delta t}{(\Delta y)^2} (1 - \cos k_2 \Delta y) \right]. \end{aligned} \quad (8.34)$$

¹⁾ См. Шортли и Уэллер [1938], Д. Янг [1950], Дж. Дуглас [1955], Писмэн и Ракфорд [1955], С. Франкел [1950], Либман [1918], Шортли [1952], Старк [1956].

²⁾ Основы таких вычислений изложены в работе Эрлиха и Гурвица [1954].

Поэтому множитель перехода для системы (8.33) равен

$$G(\Delta t, k) = \frac{1 + (1 - \theta)\psi}{1 - \theta\psi}. \quad (8.35)$$

Чтобы найти условие устойчивости, нужно знать максимум и минимум ψ как функции действительных переменных $k_1\Delta x$ и $k_2\Delta y$. Эта функция является аналитической и периодической, и, следовательно, ее экстремумы достигаются при

$$\frac{\partial \psi}{\partial (k_1 \Delta x)} = \frac{\partial \psi}{\partial (k_2 \Delta y)} = 0.$$

Непосредственный подсчет показывает, что если выполнены неравенства (8.30), то эти экстремумы могут иметь место только тогда, когда $\sin k_1\Delta x = \sin k_2\Delta y = 0$. Поэтому мы должны исследовать только четыре возможности: $\cos k_1\Delta x = \pm 1$ и $\cos k_2\Delta y = \pm 1$. Таким путем мы легко устанавливаем, что

$$-4 \left[\frac{A \Delta t}{(\Delta x)^2} + \frac{C \Delta t}{(\Delta y)^2} \right] \leq \psi \leq 0.$$

Равенство (8.35) тогда показывает, что 1) G действительно, 2) в любом случае $G \leq 1$, 3) если $\theta \geq 1/2$, то $G \geq -1$; 4) если $0 \leq \theta < 1/2$, то требование устойчивости $G \geq -1$ налагает некоторое ограничение на Δt . Таким образом, мы приходим к следующему условию устойчивости:

$$\begin{aligned} &\text{ограничений нет при } 1/2 \leq \theta \leq 1, \\ &\frac{A \Delta t}{(\Delta x)^2} + \frac{C \Delta t}{(\Delta y)^2} \leq \frac{1}{2 - 4\theta} \quad \text{при } 0 \leq \theta < 1/2. \end{aligned} \quad (8.36)$$

Это — естественное обобщение условия устойчивости, найденного для соответствующей одномерной задачи (8.4).

Для дальнейших примеров рассмотрим теперь обобщения некоторых схем повышенной точности табл. 8. I на случай двумерного уравнения диффузии

$$\frac{\partial u}{\partial t} = \sigma \left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) = \sigma \nabla^2 u. \quad (8.37)$$

Чтобы получить точное выражение¹⁾ для лапласиана, мы используем в плоскости x, y сетку, ячейками которой являются равносторонние треугольники, как указано на рис. 8.2. Шесть ближайших точек к точке (x, y) суть $(x \pm h, y)$, $(x \pm 1/2 h, y \pm 1/2 \sqrt{3} h)$, где h — расстояние между соседними точками. Если узлы сетки обозначать индексами j и l так, что $x_l = jh$, $y_l = 1/2 \sqrt{3} lh$, то ближайшими соседями точки (j, l)

¹⁾ Ср. с книгой Коллатца [1953], стр. 307.

будут точки $(j \pm 1, l)$ и $(j \pm 1/2, l \pm 1)$. Обозначим через $(\sum u)_{jl}$ сумму значений u в этих шести соседних точках, т. е.

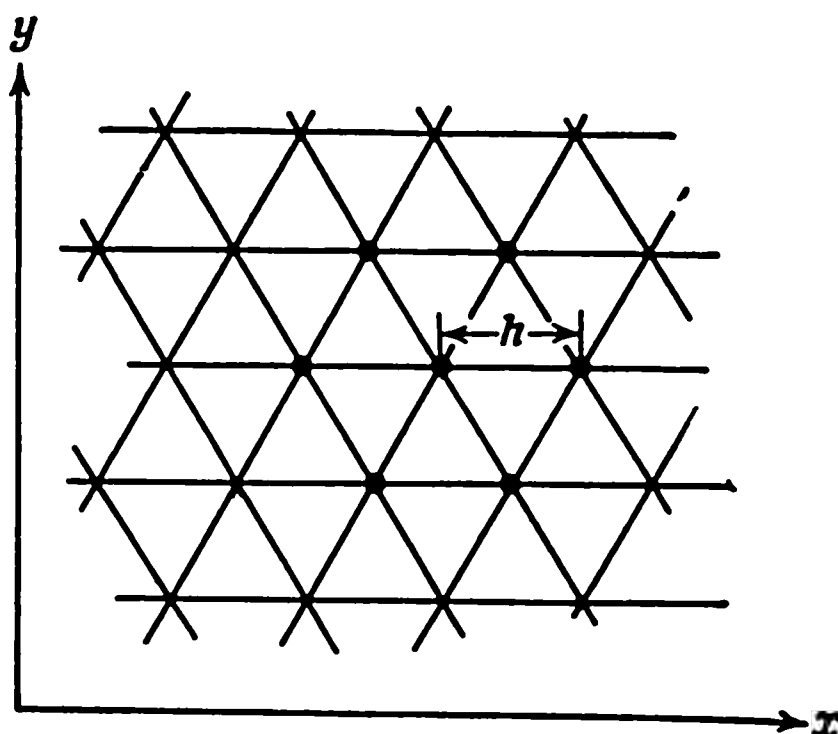
$$(\sum u)_{jl} = u_{j+1, l} + u_{j-1, l} + u_{j+1/2, l+1} + u_{j-1/2, l+1} + \\ + u_{j+1/2, l-1} + u_{j-1/2, l-1}.$$

Используя разложения в ряды Тейлора, легко усмотреть, что

$$(\sum u)_{jl} = 6u_{jl} + \frac{3}{2}h^2 (\nabla^2 u)_{jl} + \frac{3}{32}h^4 (\nabla^4 u)_{jl} + O(h^6).$$

(Если использовать четыре соседние точки прямоугольной сетки, то получится аналогичная формула, но член четвертого поряд-

Р и с. 8.2. Гексагональная сетка в плоскости (x, y) , используемая в системах (8.38)–(8.40).



ка будет $\frac{1}{12}h^4 (\partial^4 u / \partial x^4 + \partial^4 u / \partial y^4)$ и не может быть представлен через лапласиан от u). Поэтому

$$\frac{(\sum u)_{jl} - 6u_{jl}}{\frac{3}{2}h^2} = (\nabla^2 u)_{jl} + \frac{1}{16}h^2 (\nabla^4 u)_{jl} + O(h^4).$$

Следовательно, если u удовлетворяет уравнению $(\partial u / \partial t) = \sigma \nabla^2 u$, то мы имеем

$$\sigma \frac{(\sum u)_{jl} - 6u_{jl}}{\frac{3}{2}h^2} = \left(\frac{\partial u}{\partial t} \right)_{jl} + \frac{1}{16} h^2 \left(\nabla^2 \frac{\partial u}{\partial t} \right)_{jl} + O(h^4) = \\ = \left(\frac{\partial u}{\partial t} \right)_{jl} + \frac{1}{16} h^2 \frac{(\sum \frac{\partial u}{\partial t})_{jl} - 6 \left(\frac{\partial u}{\partial t} \right)_{jl}}{\frac{3}{2}h^2} + O(h^4) = \\ = \frac{3}{4} \left(\frac{\partial u}{\partial t} \right)_{jl} + \frac{1}{24} \left(\sum \frac{\partial u}{\partial t} \right)_{jl} + O(h^4).$$

Веса $\frac{3}{4}$ и $\frac{1}{24}$ соответствуют весам $\frac{5}{6}$ и $\frac{1}{12}$ в примерах 12 и 13 табл. 8.1 для одномерного уравнения диффузии.

Аналогами уравнений 12 и 13 табл. 8. I являются

$$\begin{aligned} \frac{1}{24} \frac{(\sum u)_{jl}^{n+1} - (\sum u)_{jl}^n}{\Delta t} + \frac{3}{4} \frac{u_{jl}^{n+1} - u_{jl}^n}{\Delta t} = \\ = \frac{\sigma}{2} \frac{(\sum u)_{jl}^{n+1} - 6u_{jl}^{n+1}}{3/2 h^2} + \frac{\sigma}{2} \frac{(\sum u)_{jl}^n - 6u_{jl}^n}{3/2 h^2} \end{aligned} \quad (8.38)$$

и

$$\begin{aligned} \frac{1}{24} \frac{3/2 (\sum u)_{jl}^{n+1} - 2(\sum u)_{jl}^n + 1/2 (\sum u)_{jl}^{n-1}}{\Delta t} + \\ + \frac{3}{4} \frac{3/2 u_{jl}^{n+1} - 2u_{jl}^n + 1/2 u_{jl}^{n-1}}{\Delta t} = \sigma \frac{(\sum u)_{jl}^{n+1} - 6u_{jl}^{n+1}}{3/2 h^2}. \end{aligned} \quad (8.39)$$

Мы будем называть их 14-точечной и 21-точечной формулами соответственно (в любом из этих случаев в каждой плоскости $t = \text{const}$ берется семь точек). Погрешность аппроксимации в обоих случаях равна $O[(\Delta t)^2] + O(h^4)$.

Аналогом менее точного уравнения 9 табл. 8. I является 9-точечная формула

$$\frac{3/2 u_{jl}^{n+1} - 2u_{jl}^n + 1/2 u_{jl}^{n-1}}{\Delta t} = \sigma \frac{(\sum u)_{jl}^{n+1} - 6u_{jl}^{n+1}}{3/2 h^2}, \quad (8.40)$$

для которой погрешность аппроксимации равна $O[(\Delta t)^2] + O(h^2)$.

Эти формулы согласуются с уравнением (8.37) и безусловно устойчивы при $\Delta t \rightarrow 0$ и $h \rightarrow 0$. Существует соответствие между 14-точечной формулой (8.38) и 21-точечной формулой (8.39), которое аналогично отмеченному ранее соотношению между схемами 12 и 13 табл. 8. I: первая несколько предпочтительнее для гладких начальных данных, а вторая — для быстро меняющихся или разрывных начальных данных.

Многочисленные пробные расчеты подтвердили, что эти формулы весьма точны. При этом неявные уравнения на каждом шагу по времени решались релаксационными методами. В следующем параграфе будет показано, как для этих уравнений можно использовать методы чередующихся направлений.

Следует отметить, что 21-точечная формула несколько более точна, чем 9-точечная, но, возможно, не настолько, насколько можно было бы ожидать. Нужны более точные формулы для таких приложений, как интегрирование диффузионно-возрастного уравнения для ядерного реактора в случае двух или трех пространственных переменных. Практические соображения, связанные со скоростью и объемом оперативной памяти вычислительных машин, требуют применения для таких вычислений весьма крупной сетки, и поэтому уменьшение погрешности аппроксимации является в высшей степени желательным. Для про-

стого уравнения диффузии с постоянным коэффициентом диффузии очень точное разностное уравнение может быть получено из точного решения

$$u(x, t + \Delta t) = \left(\frac{1}{4\pi\sigma \Delta t} \right)^{3/2} \int \int \int u(y, t) \exp \left[-\frac{(x-y)^2}{4\sigma \Delta t} \right] dy$$

путем аппроксимации интеграла с помощью подходящей численной квадратуры, однако обобщение такого подхода на практические задачи с переменными коэффициентами представляется весьма трудным.

§ 8.8. Методы чередующихся направлений

Метод решения многомерных задач теплопроводности и диффузии, предложенный Писмэном и Ракфордом [1955] и Дугласом [1955], сочетает безусловную устойчивость и простоту численной реализации. Сейчас уже имеются многочисленные варианты этого метода, применяемые для решения эллиптических и гиперболических задач точно так же, как и для решения систем параболических уравнений (см. работу Дугласа и Ганна [1964] и статьи, на которые они ссылаются).

Для простого уравнения диффузии (8.37) при расчете на квадратной сетке, у которой $\Delta x = \Delta y$, в последовательные периоды времени с шагом $\Delta t/2$ попеременно применяются две разностные аппроксимации:

$$u^{n+1/2} - u^n = 1/2 \alpha (\delta_x^2 u^{n+1/2} + \delta_y^2 u^n), \quad (8.41a)$$

$$u^{n+1} - u^{n+1/2} = 1/2 \alpha (\delta_x^2 u^{n+1/2} + \delta_y^2 u^{n+1}), \quad (8.41b)$$

где u^n стоит вместо u_{jl}^n и т. д., $\delta_x^2 u$ — вместо $u_{j+1,l} - 2u_{jl} + u_{j-1,l}$, $\delta_y^2 u$ — вместо $u_{j,l+1} - 2u_{jl} + u_{j,l-1}$ и α — вместо $\sigma \Delta t / (\Delta x)^2$. Предполагая для простоты, что граничные значения заданы на сторонах некоторого прямоугольника в плоскости x, y , уравнения (8.41) можно решить методом, изложенным в § 8.5: для любого фиксированного l матрица коэффициентов при неизвестных $u_{jl}^{n+1/2}$ в (8.41a) является трехдиагональной, а для фиксированного j точно такой же будет матрица коэффициентов при неизвестных u_{jl}^{n+1} в (8.41b).

Два таких последовательных временных шага образуют законченный цикл вычислений, и множитель перехода g при этом равен произведению $g'g''$ двух множителей, соответствующих каждому из уравнений. Если β_1 и β_2 обозначают $k_1 \Delta x$ и $k_2 \Delta x$, где k_1 и k_2 — частоты компоненты Фурье $\exp[i(k_1 x + k_2 y)]$, то

$$[1 + \alpha(1 - \cos \beta_1)] g' = 1 - \alpha(1 - \cos \beta_2)$$

и

$$[1 + \alpha(1 - \cos \beta_2)] g'' = 1 - \alpha(1 - \cos \beta_1),$$

так что

$$g = \frac{[1 - \alpha(1 - \cos \beta_1)][1 - \alpha(1 - \cos \beta_2)]}{[1 + \alpha(1 - \cos \beta_1)][1 + \alpha(1 - \cos \beta_2)]}; \quad (8.42)$$

поскольку $\alpha(1 - \cos \beta_{1,2}) \geq 0$, то g лежит между -1 и 1 , а это означает безусловную устойчивость схемы (8.41).

Для нахождения порядка точности вычтем (8.41б) из (8.41а) и получим

$$2u^{n+1/2} = u^{n+1} + u^n - 1/2 \alpha \delta_y^2 (u^{n+1} - u^n);$$

подстановка этого результата в сумму тех же уравнений дает

$$\frac{u^{n+1} - u^n}{\Delta t} = \sigma \frac{(\delta_x^2 + \delta_y^2)(u^{n+1} + u^n)}{2(\Delta x)^2} - \sigma^2 \Delta t \frac{\delta_x^2 \delta_y^2 (u^{n+1} - u^n)}{4(\Delta x)^4}. \quad (8.43)$$

для достаточно гладкой функции $u(x, y, t)$ последний член равен $O[(\Delta t)^2]$, а остальные члены центрированы так же, как и в уравнении Кранка — Николсона; следовательно, $e = O[(\Delta t)^2] + O[(\Delta x)^2]$.

В наиболее очевидном обобщении на случай трех переменных с помощью трех уравнений, содержащих соответственно только u^n и $u^{n+1/2}$, $u^{n+1/2}$ и $u^{n+2/2}$, $u^{n+2/2}$ и u^{n+1} и идентичных с точностью до циклической перестановки производных по x , y и z в неявном члене, безусловная устойчивость теряется и погрешность аппроксимации становится равной $O(\Delta t) + O[(\Delta x)^2]$. Поэтому был разработан другой алгоритм, который является частным случаем очень общего метода для получения схем с чередующимися направлениями из вполне неявных схем, предложенного Дугласом и Ганном [1964]. Лучше всего этот алгоритм описывается с помощью последовательности аппроксимирующих решений u^{*n+1} , u^{**n+1} и т. д. уравнения Кранка — Николсона. Последнее из них u^{***n+1} в случае трех переменных отождествляется с u^{n+1} и используется на следующем шаге по времени. Разностные уравнения имеют вид

$$\begin{aligned} u^{*n+1} - u^n &= \frac{\alpha}{2} [\delta_x^2 (u^{*n+1} + u^n) + \delta_y^2 (2u^n) + \delta_z^2 (2u^n)], \\ u^{**n+1} - u^n &= \frac{\alpha}{2} [\delta_x^2 (u^{*n+1} + u^n) + \delta_y^2 (u^{**n+1} + u^n) + \\ &\quad + \delta_z^2 (2u^n)], \quad (8.44) \\ u^{n+1} - u^n &= \frac{\alpha}{2} [\delta_x^2 (u^{*n+1} + u^n) + \delta_y^2 (u^{**n+1} + u^n) + \\ &\quad + \delta_z^2 (u^{n+1} + u^n)]. \end{aligned}$$

Очевидное обобщение оказывается эффективным при любой размерности пространства; в случае двух переменных x и y

снова получается схема (8.41), хотя это совпадение и не является очевидным.

Множитель перехода для полного шага легко вычисляется:

$$g = \frac{1 - a - b - c + ab + ac + bc + abc}{(1 + a)(1 + b)(1 + c)}, \quad (8.45)$$

где a , b и c стоят вместо $\alpha(1 - \cos k_i \Delta x)$, $i = 1, 2, 3$. Знаменатель в (8.45) после выполнения умножений содержит в точности те же члены, что и числитель, но все члены знаменателя имеют знак плюс, и поэтому снова $-1 \leq g \leq 1$, откуда следует безусловная устойчивость. Интересно отметить, что если в третьем из уравнений (8.44) в члене, соответствующем второй производной по x , заменить u^{*n+1} на u^{**n+1} (можно подумать, что аппроксимация при этом всегда будет лучше), то безусловная устойчивость снова утрачивается (g может быть меньше, чем -1).

Нахождение u_{jlk}^{n+1} из системы (8.44) снова сводится к решению последовательности трехдиагональных линейных систем. Вычисления несколько упрощаются, если вычесть второе уравнение из первого и третье из второго и затем использовать полученные уравнения вместо двух последних уравнений системы (8.44).

Погрешность аппроксимации может быть найдена путем исключения u^{*n+1} и u^{**n+1} из (8.44), которое проводится так же, как и исключение $u^{n+1/2}$ из (8.41). Окончательный результат имеет вид

$$\begin{aligned} \frac{u^{n+1} - u^n}{\Delta t} = & \sigma \frac{(\delta_x^2 + \delta_y^2 + \delta_z^2)(u^{n+1} + u^n)}{2(\Delta x)^2} - \\ & - \sigma^2(\Delta t) \frac{(\delta_x^2 \delta_y^2 + \delta_x^2 \delta_z^2 + \delta_y^2 \delta_z^2)(u^{n+1} - u^n)}{4(\Delta x)^4} + \\ & + \sigma^3(\Delta t)^2 \frac{\delta_x^2 \delta_y^2 \delta_z^2 (u^{n+1} - u^n)}{8(\Delta x)^6} \end{aligned} \quad (8.46)$$

и показывает, что в этом случае снова $e = O[(\Delta t)^2] + O[(\Delta x)^2]$.

В общей формулировке алгоритма чередующихся направлений разностное уравнение (3.4) записывается в виде

$$(I + A)u^{n+1} + Bu^n = 0, \quad (8.47)$$

где

$$A = \sum_{i=1}^q A_i$$

и каждый разностный оператор A_i (такой, как δ_x^2 , δ_y^2 и т. д.) выбирается так, чтобы оператор $I + A_i$ можно было всегда легко обратить путем решения трехдиагональной линейной системы. Обозначим последовательные приближения к u^{n+1} через

$u_{n+1}(j)$, где j соответствует числу звездочек у функций из системы (8.44). Тогда обобщением системы (8.44) будет система

$$\begin{aligned} u^{n+1}(1) + A_1 u^{n+1}(1) + \sum_{i=2}^q A_i u^n + B u^n &= 0, \\ \vdots \\ u^{n+1}(j) + \sum_{i=1}^j A_i u^{n+1}(i) + \sum_{i=j+1}^q A_i u^n + B u^n &= 0, \\ u^{n+1}(q) + \sum_{i=1}^q A_i u^{n+1}(i) + B u^n &= 0, \end{aligned} \quad (8.48)$$

в которой $u^{n+1}(q)$ отождествляется с u^{n+1} и используется на следующем шаге по времени. В j -м уравнении системы (8.48) неизвестные входят только в первый член и в член, для которого $i = j$. Дальнейшие обобщения были построены Дугласом и Ганном [1964] для слабо нелинейных уравнений (у которых нелинейным является только нелинейный член) и для многослойных схем вида, описанного в гл. 7.

Было потрачено много усилий для выяснения влияния граничных условий и переменных коэффициентов на устойчивость и точность. При очень общих предположениях можно показать, что схема (8.48) согласована с дифференциальным уравнением, если с ним согласована исходная схема (8.47). Если $q = 2$ (это ограничивает нас двумя пространственными переменными, хотя нужно заметить, что даже для двух пространственных переменных иногда необходимо выбирать $q > 2$, как это будет видно из приводимого ниже примера), то для довольно широкого класса параболических задач из устойчивости схемы (8.47) при любом фиксированном значении величины $\Delta t[(\Delta x)^{-2} + (\Delta y)^{-2}]$ следует устойчивость схемы (8.48). Если q и размерность пространства больше двух, то устойчивость схемы (8.48) следует из устойчивости схемы (8.47) при весьма частных предположениях (предполагается, что разностные операторы A_i и B коммутативны и эрмитовы, а A_i к тому же положительно определенные — это практически сводится к простой задаче теплопроводности для прямоугольника). При тех же предположениях можно также доказать, что потеря точности при переходе от (8.47) к (8.48) не превышает $O[(\Delta t)^2]$. При несколько более общих предположениях можно показать, что существует такая функция g , что схема (8.48) будет устойчивой при $\Delta t \leq g(\Delta x)$ (при этом предполагается, что $\Delta x = \Delta y$ и т. д.), или что устойчивость схемы (8.48) имеет место при достаточно малом фиксированном значении величины $\Delta t/(\Delta x)^4$. Обсуждение этих результатов и приложения изложенного метода содержатся в работе Дугласа и

Ганна [1964]. Интуиция подсказывает, что этот метод обладает устойчивостью и точностью при более широких предположениях, чем те, для которых были получены строгие доказательства.

В качестве иллюстрации применим общий алгоритм к неявному уравнению (8.38), сетка для которого изображена на рис. 8.2. Определим вторые разности по пространству δ_1^2 , δ_2^2 , δ_3^2 следующим образом:

$$\begin{aligned}\delta_1^2 u &= u(x+h, y) - 2u(x, y) + u(x-h, y), \\ \delta_{2,3}^2 u &= u(x + 1/2 h, y \pm 1/2 \sqrt{3} h) - 2u(x, y) + \\ &+ u(x - 1/2 h, y \mp 1/2 \sqrt{3} h).\end{aligned}\quad (8.49)$$

Тогда уравнение (8.38) может быть представлено в виде

$$u^{n+1} - u^n = \alpha \delta^2 (\theta u^{n+1} + (1 - \theta) u^n), \quad (8.50)$$

где $\delta^2 = \delta_1^2 + \delta_2^2 + \delta_3^2$, $\alpha = 2\sigma\Delta t/3(\Delta x)^2$, а $\theta = 1/2 - 1/24\alpha$. Теперь эти уравнения совпадают с уравнениями (8.44) за исключением того, что разностные операторы δ_x , δ_y и δ_z заменены на δ_1 , δ_2 и δ_3 , а среднее значение $1/2 u^{n+1} + 1/2 u^n$ заменено на взвешенное среднее $\theta u^{n+1} + (1 - \theta) u^n$. Иначе говоря, в (8.48) мы теперь имеем $q=3$, $A_i = -\alpha\theta\delta_i^2$ ($i=1, 2, 3$) и $B = -I\alpha(1-\theta)\delta^2$. Заметим, что q , равное трем, превосходит размерность пространства, которая равна двум.

Для нахождения множителя перехода мы используем метод, применимый в скалярном случае к системе (8.48) общего вида, если операторы A_i являются положительно определенными. (Напомним читателю, что здесь, как и всегда в рамках теории фон Неймана, мы рассматриваем краевую задачу без граничных условий и с постоянными коэффициентами. Как было отмечено Биркгофом и Варгой [1959], граничные условия, заданные на границе области, отличной от прямоугольника, приводят к некоммутативности A_i , если A_i рассматриваются как операторы, области определения которых состоят из функций, удовлетворяющих граничным условиям; таким образом, для задач с переменными коэффициентами введенный выше множитель перехода является локальным свойством системы, определенным для внутренних точек физической области, где коэффициенты меняются медленно. Иначе говоря, состояние общей теории, как уже отмечалось выше, пока что нельзя считать удовлетворительным с точки зрения практических приложений.) В результате применения оператора A_i к функции $\exp[ik_1 x + ik_2 y]$ она умножается на $2\alpha\theta(1 - \cos \beta_i)$, где

$$\beta_1 = k_1 h, \quad \beta_{2,3} = 1/2 k_1 h \pm 1/2 \sqrt{3} k_2 h;$$

так как этот множитель неотрицателен, оператор A_j является положительно определенным. Промежуточные величины $u^{n+1}(j)$ ($j = 1, \dots, q-1$) могут быть исключены из системы (8.48). В результате для $q = 3$ получим

$$(I + A)u^{n+1} + Bu^n + (A_1A_2 + A_2A_3 + A_1A_3 + A_1A_2A_3)(u^{n+1} - u^n) = 0. \quad (8.51)$$

Если в этом уравнении обозначить через $a = a(k)$, $b = b(k)$ и $\Delta = \Delta(k)$ множители, на которые будет умножаться $\exp(ik \cdot x)$ при применении операторов A , B и $(A_1A_2 + \dots + A_1A_2A_3)$, то множитель перехода запишется в виде

$$g = (-b + \Delta)/(1 + a + \Delta). \quad (8.52)$$

В силу (8.47) множитель перехода исходной схемы — в нашем случае схемы (8.38) — равен $g_0 = -b/(1 + a)$. Так как $\Delta \geq 0$, из условия $-1 \leq g_0 \leq 1$ следует, что $-1 \leq g \leq 1$; следовательно, метод чередующихся направлений для схемы (8.38) всегда устойчив.

§ 8.9. Методы расщепления и дробных шагов

Методы, тесно связанные (а в некоторых случаях и совпадающие) с методами чередующихся направлений, были примерно в то же время разработаны советскими математиками¹⁾ и известны теперь, как методы расщепления или методы дробных шагов. Основная идея, выдвинутая Багриновским и Годуновым [1957] и обобщенная затем Яненко [1964], состоит в следующем: пусть для решения уравнения $\partial u / \partial t = Au$, где $A = A_1 + \dots + A_q$, выбрана некоторая разностная схема; тогда оператор A в ней последовательно заменяется на $qA_1, qA_2, \dots, qA_q$, причем продвижение по времени при каждой такой замене происходит на часть $\Delta t/q$ временного шага Δt . Таким образом, можно заменить сложную (например, многомерную) задачу последовательностью более простых (например, одномерных) задач.

Например, схема Кранка — Николсона для многомерного уравнения теплопроводности

$$\begin{aligned} \frac{u^{n+1} - u^n}{\Delta t} &= \sigma \frac{(\delta_x^2 + \delta_y^2 + \dots + \delta_w^2)(u^{n+1} + u^n)}{2(\Delta x)^2} = \\ &= -(A_1 + \dots + A_q) \left(\frac{u^{n+1} + u^n}{2} \right), \end{aligned}$$

¹⁾ Багриновский и Годунов [1957], Яненко [1959], Дьяконов [1962, 1964], Софронов [1962], Коновалов [1964], Самарский [1964], Марчук и Яненко [1966]; обширная библиография приводится в последней из этих работ и в книге Яненко [1967].

где $A_i = -(\sigma/(\Delta x)^2 \delta_x^2$ и т. д., $\Delta x = \Delta y = \dots$, а δ_x^2, δ_y^2 и т. д. обозначают, как обычно, вторые разности по соответствующему переменному, заменится последовательностью уравнений

$$\frac{u^{n+i/q} - u^{n+(i-1)/q}}{\Delta t} = -A_i \frac{u^{n+i/q} + u^{n+(i-1)/q}}{2} \quad (i = 1, 2, \dots, q). \quad (8.53)$$

Отличие от схемы Дугласа (8.48) состоит в том, что в разностном операторе члены A_j при $j \neq i$ полностью опущены. Поэтому в данном случае, как и для любого скалярного уравнения, множитель перехода просто равен произведению множителей перехода для одномерных уравнений (8.53). Так как они являются уравнениями типа Кранка — Николсона, это означает безусловную устойчивость схемы (8.53). Чтобы найти погрешность аппроксимации, заметим, что

$$\left(I + \frac{\Delta t}{2} A_i\right) u^{n+i/q} = \left(I - \frac{\Delta t}{2} A_i\right) u^{n+(i-1)/q};$$

так как в нашем примере операторы A_i коммутативны, мы имеем

$$\left(I + \frac{\Delta t}{2} A_1\right) \dots \left(I + \frac{\Delta t}{2} A_q\right) u^{n+1} = \left(I - \frac{\Delta t}{2} A_1\right) \dots \left(I - \frac{\Delta t}{2} A_q\right) u^n.$$

После выполнения умножений получим

$$\begin{aligned} u^{n+1} - u^n + \Delta t A \frac{u^{n+1} + u^n}{2} = \\ = -\frac{(\Delta t)^2}{4} (A_1 A_2 + A_1 A_3 + \dots + A_{q-1} A_q) (u^{n+1} - u^n) + O[(\Delta t)^3] \end{aligned}$$

так как $u^{n+1} - u^n = O(\Delta t)$, а левая часть представляет схему Кранка — Николсона, то мы видим (после деления на Δt), что погрешность аппроксимации равна $O[(\Delta t)^2] + O[(\Delta x)^2]$.

Глава 9

УРАВНЕНИЕ ПЕРЕНОСА

§ 9.1. Физические основы

Задачи рассматриваемого в этой главе типа возникают главным образом в двух областях: диффузия нейтронов и перенос лучистой энергии. Термин «диффузия нейтронов» часто (и неудачно) применяется для описания миграции и размножения нейтронов даже в том случае, когда приближение диффузии неприменимо. Основное интегро-дифференциальное уравнение для этих процессов называется уравнением переноса. Иногда используется неправильный термин «уравнение Больцмана».

Уравнение переноса описывает движение частиц (нейтронов, протонов и т. п.), которые между соударениями перемещаются прямолинейно с постоянной скоростью, но при прохождении через вещество постоянно испытывают с определенной вероятностью соударения с ядрами или атомами и отклоняются от первоначального направления, замедляются, поглощаются или размножаются. Уравнение описывает настолько большой ансамбль частиц, что его можно рассматривать как континуум, пренебрегая статистическими флуктуациями. В уравнении Больцмана из статистической механики рассеивающие частицы таковы же, как и рассеиваемые, в то время как в уравнении переноса распределение и свойства рассеивающих центров предполагаются заданными; поэтому уравнение переноса является линейным, тогда как уравнение Больцмана квадратично.

В любой момент времени движение нейтрона может быть представлено точкой шестимерного фазового пространства, координаты которой являются компонентами радиуса-вектора и вектора скорости нейтрона. Весь ансамбль в целом описывается своим распределением в этом пространстве. Краевая задача состоит в том, чтобы найти изменение этого распределения со временем, если в начальный момент оно задано.

Аналогичное положение имеет место и в задачах о переносе лучистой энергии с той лишь разницей, что координатами фазового пространства являются компоненты радиуса-вектора и вектора *импульса* (это можно было бы сделать также и в случае нейтронов), так как скорость фотона в этих задачах не зависит от его энергии $h\nu$. В астрофизике основной интерес представ-

ляют задачи, связанные с равновесным состоянием, однако для пульсирующих и иных переменных звезд возникают задачи, учитывающие зависимость от времени.

§ 9.2. Общее уравнение переноса нейтронов

Предположим, что система, в которой происходят нейтронные реакции, находится в выпуклой области $\mathcal{R}$, ограниченной поверхностью $\mathcal{S}$. Пусть $\mathbf{x} = (x, y, z)$ и $\mathbf{v} = (v_x, v_y, v_z)$ обозначают радиус-вектор и вектор скорости, и пусть для $\mathbf{x} \in \mathcal{R}$

$$\psi(\mathbf{x}, \mathbf{v}, t) d\mathbf{x} d\mathbf{v}$$

— число нейтронов в интервале $(x, x + dx)$, $(y, y + dy)$, $\dots$, $(v_z, v_z + dv_z)$ в момент времени t , причем $d\mathbf{x}$ — краткое обозначение $dx dy dz$, а $d\mathbf{v}$ аналогично означает $dv_x dv_y dv_z$.

Предположим (как это делается в обычной статистической механике), что $d\mathbf{x}$ и $d\mathbf{v}$ можно выбрать так, чтобы они были малы по сравнению с областями, в которых условия заметно меняются, и в то же время достаточно велики для того, чтобы $\psi d\mathbf{x} d\mathbf{v}$ представляло большое число нейтронов. Флуктуационные явления нас здесь не интересуют.

Для наших целей достаточно предположить, что нейтроны являются классическими точечными частицами, соударения которых с ядрами происходят по статистическим законам, так что мы будем иметь дело со средними значениями и вероятностями. Пусть $v\sigma(\mathbf{v}, \mathbf{x})$ означает вероятность столкновения за единицу времени нейтрона, обладающего скоростью $\mathbf{v}$ (v — абсолютное значение $\mathbf{v}$). Величина $\sigma(\mathbf{v}, \mathbf{x})$ является полным поперечным сечением ядра, отнесенным к единице объема в точке $\mathbf{x}$. Если вещество изотропно, то σ не зависит от направления $\mathbf{v}$, а если оно однородно, то σ не зависит от $\mathbf{x}$; поэтому вместо $\sigma(\mathbf{v}, \mathbf{x})$ мы будем писать $\sigma(v)$, хотя во многих интересных задачах σ зависит от $\mathbf{x}$ по крайней мере в том смысле, что система состоит из нескольких областей, в которых $\sigma(v)$ имеет разные значения.

Если соударяющийся нейтрон имеет скорость $\mathbf{v}'$, то через $K(\mathbf{v}', \mathbf{v}) d\mathbf{v}$ мы будем обозначать вероятность того, что он после соударения будет иметь компоненты скорости в интервалах $(v_x, v_x + dv_x)$, $(v_y, v_y + dv_y)$, $(v_z, v_z + dv_z)$. В случае изотропной среды K может зависеть от скоростей $\mathbf{v}'$ и $\mathbf{v}$ и от угла между ними, но не от самих направлений $\mathbf{v}'$ и $\mathbf{v}$. При заданных $\mathbf{v}'$ и $\mathbf{x}$ интеграл $\int K d\mathbf{v}$, взятый по всем скоростям испускаемого в процессе соударения нейтрона, дает среднее число нейтронов, испускаемых при соударении. Обозначим этот интеграл через $1 + f$; тогда при $f = 0$ среда только рассеивает нейтроны, при $f < 0$ она может их поглощать, а $f > 0$ соответствует среде,

в которой нейтроны могут размножаться. В полевом случае рассматривают такие процессы, как деление и $(n, 2n)$ -реакция, следствием которых является испускание двух и более нейтронов. В этом случае слово «он» в данном выше определении K должно быть заменено словом «нейтрон», который не обязательно окажется начальным нейтроном.

Поведение $\psi(x, v, t)$ описывается *уравнением переноса*

$$(\partial/\partial t + v \cdot \nabla) \psi(x, v, t) = -v\sigma(v) \psi(x, v, t) + \int v'\sigma(v') K(v', v) \psi(x, v', t) dv', \quad (9.1)$$

где $v \cdot \nabla$ означает $v_x \partial/\partial x + v_y \partial/\partial y + v_z \partial/\partial z$, а $\int \dots dv'$ означает

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots dv'_x dv'_y dv'_z.$$

Граничное условие заключается в том, что нейтроны извне не попадают в систему. Предположим, что область $\mathcal{R}$ выпукла (иначе мы должны были бы учитывать нейтроны, вылетающие из системы в одном месте и возвращающиеся в нее в другом) и что поверхность $\mathcal{S}$ почти везде имеет нормаль. Если эти предположения не выполняются в данной области $\mathcal{R}$, то можно заменить ее большей областью, удовлетворяющей поставленным требованиям, положив при этом в дополнительно присоединенном пространстве сечение равным нулю. В качестве такой новой области $\mathcal{R}$ мы можем выбрать сферу. Если $n(x)$ — внешняя по отношению к $\mathcal{S}$ нормаль в точке x , то граничное условие примет вид

$$\psi(x, v, t) = 0 \text{ при } x \in \mathcal{S} \text{ и } v \cdot n(x) < 0. \quad (9.2)$$

Интерпретация уравнения переноса (или, вернее, физическая картина, на основании которой оно обычно выводится) следующая: ансамбль нейтронов можно представлять себе как суперпозицию очень большого числа (в действительности пятипараметрического семейства) узких пучков движущихся вместе нейтронов. Функция $\psi(x, v, t)$ представляет интенсивность в точке x одного из таких пучков, а левая часть уравнения (9.1) — изменение этой интенсивности в зависимости от времени, которое мы будем наблюдать, двигаясь вместе с частицей. Первый член правой части описывает выбывание частиц из пучка вследствие соударений, а последний член (интеграл) соответствует приходу частиц в рассматриваемый пучок вследствие других соударений.

Если функция $\psi(x, v, 0)$ задана, то уравнения (9.1) и (9.2) приводят к линейной краевой задаче. Можно написать $\partial\psi/\partial t =$

$= A\psi$, где A — линейный оператор, который дается уравнением (9.1) и область определения которого состоит из функций, удовлетворяющих граничному условию (9.2).

Так как число независимых переменных велико, то мы не можем применять численные методы непосредственно к полному уравнению переноса. Вместо этого, сделав упрощающие предположения, рассмотрим несколько частных случаев. Мы уже предположили, что $\sigma(v)$ и $K(v', v)$ не зависят от времени, а наши обозначения соответствуют предположению изотропности и однородности вещества. Можно сделать дополнительные предположения (не все одновременно):

- 1) число групп нейтронов конечно,
- 2) имеет место симметрия слоя,
- 3) имеет место сферическая симметрия,
- 4) имеет место цилиндрическая симметрия,
- 5) имеет место осевая симметрия,
- 6) зависимость от координат отсутствует,
- 7) $1/\sigma$ очень мало,
- 8) испускаемые при столкновении нейтроны распределены изотропно.

Предположение 1) означает, что скорость v не меняется непрерывно, а принимает конечное число значений $v_1, v_2, \dots, v_G$ и что интегрирование по v заменяется суммированием по группам нейтронов. В частности, если предположить, что все нейтроны деления имеют, грубо говоря, одну и ту же энергию (и, следовательно, одну и ту же скорость) и что при рассеянии энергия не меняется, то имеется только одна группа. Предположение 2) означает, во-первых, что все величины не зависят от двух декартовых координат, скажем x и y , и, во-вторых, что распределение скоростей симметрично по отношению к направлению z ; другими словами, система и ансамбль нейтронов в ней инвариантны по отношению к переносам и вращениям в плоскости x, y . Предположение 3) означает инвариантность по отношению к вращениям вокруг точки; предположение 4) — инвариантность по отношению к переносам параллельно оси; предположение 5) — инвариантность относительно вращений вокруг оси; предположение 6) — инвариантность относительно всех переносов и вращений в пространстве. Предположение 7) означает, что средний свободный пробег $1/\sigma$ мал по сравнению со всеми характерными длинами, относящимися к системе или к распределению в ней нейтронов (это соответствует предельному случаю теории диффузии). Наконец, предположение 8) означает, что $K(v', v)$ не зависит от направления v , а потому и от угла между v' и v .

Мы не будем рассматривать все возможные комбинации этих предположений, а ограничимся лишь несколькими типичными случаями.

§ 9.3. Однородный слой; одна группа

Все нейтроны имеют скорость v , а слой заполняет область $-a \leq z \leq a$. Вероятность того, что при рассеянии направление вылета будет лежать в единице телесного угла и будет составлять угол α с первоначальным направлением, обозначается через $\frac{1}{4\pi} p(\cos \alpha)$ — это выражение подставляется вместо ядра $K(v', v)$ в общее уравнение; случай изотропии соответствует равенству

$$p(\cos \alpha) = 1.$$

Пусть θ — угол между вектором скорости v и осью z , и пусть $\mu = \cos \theta$; тогда функцию распределения можно записать в виде $\psi(z, \mu, t)$; такая запись соответствует числу нейтронов, отнесенному к единице объема пространства и единице телесного угла в направлении движения. После деления на v уравнение переноса запишется в виде

$$\begin{aligned} \left(\frac{1}{v} \frac{\partial}{\partial t} + \mu \frac{\partial}{\partial z} + \sigma \right) \psi(z, \mu, t) = \\ = \sigma \frac{1+f}{4\pi} \int_{-1}^1 d\mu' \int_0^{2\pi} p(\cos \alpha) \psi(z, \mu', t) d\varphi \\ \text{при } -a \leq z \leq a, \quad t \geq 0, \end{aligned} \quad (9.3)$$

где

$$\cos \alpha = \mu\mu' + \sqrt{(1-\mu^2)(1-\mu'^2)} \cos \varphi. \quad (9.4)$$

Здесь α — угол между направлением скорости нейтрона до столкновения и после столкновения; полярные углы этих направлений можно записать как $(\theta, 0)$ и (θ', φ) , где $\mu = \cos \theta$ и $\mu' = \cos \theta'$.

Граничные условия выглядят так:

$$\psi(z, \mu, t) = 0 \text{ при } z = a, \mu < 0 \text{ и при } z = -a, \mu > 0. \quad (9.5)$$

Это соответствует тому, что нейтроны не попадают на левую сторону слоя слева и на правую сторону — справа.

Во многих случаях угловое распределение при рассеянии изотропно. Это означает отсутствие корреляции между начальным и конечным направлениями, и поэтому $p(\cos \alpha) = 1$. Такое утверждение справедливо, например, при рассеянии нейтро-

нов малой энергии на тяжелых атомах. В этом случае имеем следующее уравнение переноса:

$$\left(\frac{1}{v} \frac{\partial}{\partial t} + \mu \frac{\partial}{\partial z} + \sigma\right) \psi(z, \mu, t) = \frac{\sigma(1+f)}{2} \int_{-1}^1 \psi(z, \mu', t) d\mu'. \quad (9.6)$$

Заметим, что изотропность рассеяния не приводит к изотропности решения: ψ зависит от μ , даже если $p(\cos \alpha)$ постоянно.

§ 9.4. Однородная сфера; одна группа

Пусть a — радиус сферы, r — полярный радиус, $\mu = \cos \theta$, где θ — угол между радиусом-вектором, проведенным к нейтрону из начала координат, и вектором скорости нейтрона. Уравнение переноса имеет вид

$$\begin{aligned} \left(\frac{1}{v} \frac{\partial}{\partial t} + \mu \frac{\partial}{\partial r} + \frac{1-\mu^2}{r} \frac{\partial}{\partial \mu} + \sigma\right) \psi(r, \mu, t) = \\ = \sigma \frac{1+f}{4\pi} \int_{-1}^1 d\mu' \int_0^{2\pi} p(\cos \alpha) \psi(r, \mu', t) d\varphi \\ \text{при } 0 < r \leq a, \quad t \geq 0, \end{aligned} \quad (9.7)$$

где $\cos \alpha$ определен, как и ранее, равенством (9.4); граничные условия таковы:

$$\psi(r, \mu, t) = 0 \quad \text{при } r = a, \mu < 0. \quad (9.8)$$

Можно было бы подумать, что необходимо задать граничное условие и в точке $r = 0$, как это часто делается при решении физических задач в полярных координатах, так как эта система координат имеет в данной точке особенность. Однако в нашем случае дело обстоит не так. Можно показать, что даже без такого условия решение является единственным, если оно вообще существует как решение в пространстве L_2 , за исключением прямой $\mu = 1$. Но эта прямая в пространстве скоростей имеет меру нуль, а две функции распределения, которые совпадают с точностью до множества меры нуль, представляют одно и то же физическое распределение. При этих рассуждениях предполагается, что ψ принимает конечные значения, и поэтому дельта-функция и ей подобные из рассмотрения исключаются.

§ 9.5. Метод сферических гармоник

Метод, который носит это название (см. работу Чандрасекара [1944а]), мало применялся при решении задач, содержащих зависимость от времени, но так как он успешно использовался

в стационарных задачах нейтронной физики, мы рассмотрим ряд простых систем разностных уравнений для зависящих от времени уравнений в сферических гармониках и обсудим их устойчивость. Мы ограничимся задачами со сферической симметрией и симметрией слоя (случаем одной группы), хотя обобщение на случай многих групп и осевой или цилиндрической симметрии, как и в стационарных задачах, можно сделать непосредственно.

В этом методе угловая зависимость распределения в данной точке пространства представлена разложением¹⁾ функции $\psi(r, \mu, t)$ по полиномам Лежандра от μ с коэффициентами, зависящими от r и t . Подстановка этого разложения в уравнение переноса приводит к системам уравнений в частных производных для коэффициентов. В частности, положим

$$\psi(r, \mu, t) = \sum_{l=0}^{\infty} \sqrt{2l+1} \psi_l(r, t) P_l(\mu) \quad (9.9)$$

и

$$p(\cos \alpha) = \sum_{l=0}^{\infty} p_l P_l(\cos \alpha). \quad (9.10)$$

Множитель $\sqrt{2l+1}$, входящий в разложение (9.9), носит совершенно произвольный характер — его вводят для удобства нормировки $\psi_l(r, t)$; коэффициенты p_l в формуле для рассеяния (9.10) предполагаются известными (на основании физических измерений или из других источников) — это эквивалентно предположению, что функция $p(\cos \alpha)$ известна.

Если $\cos \alpha$ определяется равенством (9.4), то, согласно теореме сложения полиномов Лежандра²⁾, мы имеем

$$P_l(\cos \alpha) = P_l(\mu) P_l(\mu') + 2 \sum_{k=1}^l P_l^k(\mu) P_l^k(\mu') \frac{(l-k)!}{(l+k)!} \cos k\varphi, \quad (9.11)$$

где P_l^k — присоединенная функция Лежандра. Если это выражение для $P_l(\cos \alpha)$ подставить в разложение (9.10) для $p(\cos \alpha)$ и проинтегрировать по φ , то мы увидим, что все члены, входящие в сумму (9.11), взаимно уничтожатся, и останется только произведение $P_l(\mu) P_l(\mu')$. Таким образом,

$$\int_0^{2\pi} p(\cos \alpha) d\varphi = 2\pi \sum_{l=0}^{\infty} p_l P_l(\mu) P_l(\mu'). \quad (9.12)$$

¹⁾ См. Курант и Гильберт [1951], т. 1, стр. 423. Во время войны это разложение было использовано для уравнения переноса несколькими исследователями. Первая опубликованная работа принадлежит Чандрасекхару [1944а].

²⁾ См. Уиттекер и Ватсон [1927], т. 2, стр. 145.

Согласно уравнению (9.7), этот результат следует умножить на разложение $\psi(r, \mu', t)$, которое можно записать в следующем виде (ср. с разложением (9.9)):

$$\psi(r, \mu', t) = \sum_{j=0}^{\infty} \sqrt{2j+1} \psi_j(r, t) P_j(\mu'),$$

и проинтегрировать по μ' от -1 до $+1$.

Так как полиномы Лежандра удовлетворяют условиям ортогональности и нормировки ¹⁾

$$\int_{-1}^1 P_l(\mu') P_j(\mu') d\mu' = \begin{cases} 0 & \text{при } l \neq j, \\ \frac{2}{2j+1} & \text{при } l = j, \end{cases}$$

результат можно упростить. Следовательно, правая часть уравнения (9.7) приводится к виду

$$\sigma(1+f) \sum_{j=0}^{\infty} (p_j \psi_j(r, t) P_j(\mu) / \sqrt{2j+1}). \quad (9.13)$$

Теперь желательно представить левую часть уравнения (9.7) также в виде суммы членов, каждый из которых представляет собой произведение полинома Лежандра от μ на функцию только от r и t — тогда мы сможем приравнять коэффициенты при $P_j(\mu)$ и получить систему уравнений для $\psi_j(r, t)$. Чтобы проделать это, мы используем разложение $\psi(r, \mu, t)$, однако мы должны избавиться от входящих в уравнение (9.7) членов, содержащих μ и $(1 - \mu^2) \partial / \partial \mu$. Этого можно добиться, применяя некоторые рекуррентные соотношения ²⁾. Одно из них имеет вид

$$\mu P_l(\mu) = \frac{(l+1) P_{l+1}(\mu) + l P_{l-1}(\mu)}{2l+1}.$$

Используя его, получаем

$$\begin{aligned} \mu \psi(r, \mu, t) &= \sum_{l=0}^{\infty} \psi_l(r, t) \frac{(l+1) P_{l+1}(\mu) + l P_{l-1}(\mu)}{\sqrt{2l+1}} = \\ &= \sum_{j=0}^{\infty} \left[\frac{j}{\sqrt{2j-1}} \psi_{j-1}(r, t) + \frac{j+1}{\sqrt{2j+3}} \psi_{j+1}(r, t) \right] P_j(\mu). \end{aligned} \quad (9.14)$$

Последнее выражение получается в результате перегруппировки членов суммы (предположим, что это возможно): в первом

¹⁾ См. Уиттекер и Ватсон [1927], т. 2, стр. 113.

²⁾ Эти соотношения подобны тем, которые указаны в книге Уиттекера и Ватсона [1927], т. 2, стр. 116, и могут быть выведены из них.

члене j равно $l + 1$, а во втором равно $l - 1$; определяем, кроме того, $\psi_{-1}(r, t)$ так, чтобы оно было равно нулю.

Второе рекуррентное соотношение записывается так:

$$(1 - \mu^2) \frac{d}{d\mu} P_l(\mu) = \frac{l(l+1)}{2l+1} [P_{l-1}(\mu) - P_{l+1}(\mu)];$$

использовав его аналогичным образом, получим

$$(1 - \mu^2) \frac{\partial}{\partial \mu} \psi(r, \mu, t) =$$

$$= \sum_{j=0}^{\infty} \left[-\frac{j(j-1)}{\sqrt{2j-1}} \psi_{j-1}(r, t) + \frac{(j+1)(j+2)}{\sqrt{2j+3}} \psi_{j+1}(r, t) \right] \times P_j(\mu). \quad (9.15)$$

Подставим теперь разложения (9.13) — (9.15) в уравнение переноса (9.7) и приравняем коэффициенты при $P_j(\mu)$ в правой и левой частях. В результате для $\psi_j(r, t)$ получится система дифференциальных уравнений в частных производных. Уравнение, в которое входит производная от ψ_j по времени, в силу разложения (9.14) и (9.15) содержит не только ψ_j , но также ψ_{j+1} и ψ_{j-1} . Поэтому если уравнения записать последовательно, то каждое из них будет связано с предыдущим и последующим.

Приближение по сферическим гармоникам порядка L является результатом решения первых $L + 1$ уравнений для $L + 1$ неизвестных функций $\psi_0, \psi_1, \dots, \psi_L$, причем в последнем из этих уравнений ψ_{L+1} полагается равной нулю.

Многое в предыдущем изложении носит эвристический характер. Необходимо обосновать использование соответствующих разложений, а также доказать допустимость перегруппировки членов рядов и справедливость предположения о том, что приближение в сферических гармониках дает решение уравнения переноса при $L \rightarrow \infty$. Б. Карлсон [1946] дал это доказательство для некоторых случаев, выразив приближение в сферических гармониках аналитически и изучив предельный переход при $L \rightarrow \infty$. Позднее Келлер и Вендрофф [1957] доказали сходимость так называемых дискретно-ординатных методов. Поскольку один из них, а именно метод Вика — Чандрасекхара, эквивалентен между сферических гармоник, это означает сходимость последнего (см. § 9.10 и 9.11). Однако мы не считаем своей задачей исследование уравнений в сферических гармониках со строгостью, большей, чем строгость, с которой было выведено уравнение переноса, служащее для них основой. Поэтому, считая краевую задачу для этих уравнений корректно поставленной, мы рассмотрим ее приближенное решение с помощью численных методов.

Чтобы записать уравнения в сферических гармониках в компактном виде, положим, что $\mathbf{u}$ означает $(L + 1)$ -мерный вектор-

столбец, определяемый следующим образом:

$$\mathbf{u} = \mathbf{u}(r, t) = \begin{pmatrix} \psi_0(r, t) \\ \psi_1(r, t) \\ \vdots \\ \psi_L(r, t) \end{pmatrix}. \quad (9.16)$$

Тогда уравнения примут вид

$$\left(\frac{1}{v} \frac{\partial}{\partial t} + M \frac{\partial}{\partial r} + \frac{1}{r} N + \sigma \right) \mathbf{u} = \sigma (1 + f) S \mathbf{u},$$

где M , N и S — квадратные матрицы порядка $L + 1$, а именно

$$M = \begin{pmatrix} 0 & \frac{1}{\sqrt{3}} & 0 & 0 & 0 & \dots \\ \frac{1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{3 \cdot 5}} & 0 & 0 & \dots \\ 0 & \frac{2}{\sqrt{3 \cdot 5}} & 0 & \frac{3}{\sqrt{5 \cdot 7}} & 0 & \dots \\ 0 & 0 & \frac{3}{\sqrt{5 \cdot 7}} & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

$$N = \begin{pmatrix} 0 & \frac{1 \cdot 2}{\sqrt{1 \cdot 3}} & 0 & 0 & \dots \\ 0 & 0 & \frac{2 \cdot 3}{\sqrt{3 \cdot 5}} & 0 & \dots \\ 0 & -\frac{1 \cdot 2}{\sqrt{3 \cdot 5}} & 0 & \frac{3 \cdot 4}{\sqrt{5 \cdot 7}} & \dots \\ 0 & 0 & -\frac{2 \cdot 3}{\sqrt{5 \cdot 7}} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

и

$$S = \begin{pmatrix} p_0 & 0 & 0 & 0 & \dots \\ 0 & \frac{p_1}{3} & 0 & 0 & \dots \\ 0 & 0 & \frac{p_2}{5} & 0 & \dots \\ 0 & 0 & 0 & \frac{p_3}{7} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Из уравнений (9.16) можно получить соответствующие уравнения для слоя, заменив r на z и положив N равным нулю.

§ 9.6. Слой; система разностных уравнений I для гиперболических уравнений

Уравнения в сферических гармониках для слоя можно записать в векторно-матричной форме

$$\left(\frac{1}{v} \frac{\partial}{\partial t} + M \frac{\partial}{\partial z} + \sigma\right) \mathbf{u} = \sigma(1 + f) S \mathbf{u}, \quad (9.17)$$

где M и S — квадратные матрицы порядка $L + 1$, определенные в предыдущем параграфе, а $\mathbf{u} = \mathbf{u}(z, t)$ — вектор с $L + 1$ компонентами. Поэтому мы имеем дело с системой $L + 1$ совместных дифференциальных уравнений в частных производных первого порядка с независимыми переменными z и t . Уравнения линейны и имеют постоянные коэффициенты, так что для исследования устойчивости и других свойств соответствующих разностных уравнений можно использовать методы, описанные в гл. 4.

Эти уравнения образуют симметричную гиперболическую систему; результаты, которые будут получены в этом и в следующих параграфах, применимы к любой такой системе.

До сих пор граничные условия для уравнений в сферических гармониках не были даны. В случае конечного слоя настоящие граничные условия (9.5) заменяются $L + 1$ условием, налагаемым на $L + 1$ функций $\psi_l(r, t)$; половина этих условий задается при $z = a$, а остальные — при $z = -a$. Ниже (в § 9.12) будет показано, как можно получить такие условия, однако их природа такова, что их нельзя заменить условием периодичности, позволяющим при анализе применять ряды Фурье. В нашем исследовании мы поэтому ограничимся случаем бесконечного слоя, в котором плотность нейтронов является интегрируемой с квадратом функцией от z , позволяющей использовать интегралы Фурье. Таким был бы, например, случай, если бы тонкий плоский источник нейтронов, помещенный в однородную среду, испускал нейтроны в течение малого времени и мы интересовались бы их последующим движением.

Система конечно-разностных уравнений, которую мы будем называть «системой I», имеет вид

$$\frac{u_j^{n+1} - u_j^n}{v \Delta t} + M \frac{u_{j+1}^n - u_{j-1}^n}{2 \Delta z} + \sigma u_j^n = \sigma(1 + f) S u_j^n. \quad (9.18)$$

При этом используется разность по времени вперед и центрированная разность по пространству. Уравнения записаны в явной форме, а погрешность аппроксимации e равна $O(\Delta t) + O((\Delta z)^2)$. Матрица перехода для коэффициентов Фурье при

e^{ikz} такова:

$$G(\Delta t, k) = (1 - v\sigma \Delta t) I + v\sigma(1 + f) \Delta t S - i \frac{v \Delta t}{\Delta z} [\sin k \Delta z] M, \quad (9.19)$$

где I — единичная матрица.

При рассмотрении вопроса об устойчивости предполагается заданным поведение отношения величин Δt и Δz при их стремлении к нулю. Мы рассмотрим два случая.

Случай 1. Пусть $\alpha_0 = v\Delta t/\Delta z$ остается постоянным при $\Delta t, \Delta z \rightarrow 0$. Легко видеть, что уравнения неустойчивы при любом значении α_0 . Чтобы убедиться в этом, достаточно рассмотреть выражение

$$G(0, k) = I - i\alpha_0 [\sin k \Delta z] M. \quad (9.20)$$

Матрица M симметрична и имеет действительные собственные значения, поэтому модуль наибольшего собственного значения $G(0, k)$ равен

$$\sqrt{1 + \alpha_0^2 \mu_0^2 \sin^2 k \Delta z}$$

(здесь μ_0 — наибольшее по абсолютной величине собственное значение M). Максимальное его значение при всех действительных k составляет

$$\sqrt{1 + \alpha_0^2 \mu_0^2} > 1$$

и не зависит от Δt , поэтому условие фон Неймана (4.18) не выполняется. Таким образом, система I неустойчива при любом фиксированном $\Delta t/\Delta z$ при $\Delta t, \Delta z \rightarrow 0$.

Случай 2. Пусть $\alpha_1 = v\Delta t/[\sigma(\Delta z)^2]$ фиксировано при $\Delta t, \Delta z \rightarrow 0$. Покажем, что система I устойчива при любом фиксированном α_1 . Для этой цели используем сокращенную запись

$$(1 + f) S - I = F, \quad v\sigma \Delta t = \varepsilon;$$

тогда

$$\begin{aligned} G &= I + \varepsilon F - i \sqrt{\varepsilon \alpha_1} [\sin k \Delta z] M, \\ G^* G &= I + 2\varepsilon F + \varepsilon^2 F^2 + \varepsilon \alpha_1 [\sin^2 k \Delta z] M^2 + \\ &\quad + i \sqrt{\alpha_1} \varepsilon^{3/2} (MF - FM) \sin k \Delta z, \end{aligned}$$

где G — матрица перехода $G(\Delta t, k)$. Заметим, что одно из произведений матриц в выражении $G^* G$ выпадает вследствие того, что единичная матрица I перестановочна с любой матрицей и, в частности, с матрицей M . Следовательно, в произведении $G^* G$ не будет члена, содержащего просто $\sqrt{\varepsilon}$, и поэтому $G^* G$ равно I плюс матрица, каждый элемент которой есть $O(\varepsilon)$. Постоянные, входящие в эти выражения $O(\varepsilon)$, являются ограниченными

функциями от k , так как k входит в G^*G только через $\sin k\Delta z$. Поэтому для всех действительных k норма¹⁾ G не превышает $1 + O(\varepsilon)$ и, следовательно, не превышает $1 + O(\Delta t)$. Итак, согласно выводам § 11 гл. 4, система I устойчива при любом фиксированном $\Delta t/(\Delta z)^2$ при $\Delta t, \Delta z \rightarrow 0$.

§ 9.7. Парадокс

С точки зрения программиста, выполняющего расчет на вычислительной машине, результаты, касающиеся системы I, кажутся парадоксальными. Конечно, он не устремит Δt и Δz к нулю, а просто использует наименьшие значения, которые соответствуют емкости и быстродействию машины и его собственным возможностям. Когда он спрашивает: «Устойчивы ли мои уравнения?», то можно только ответить, что они неустойчивы, если он рассматривает свои вычисления как часть бесконечной последовательности вычислений при фиксированном значении $\Delta t/\Delta z$, и что они устойчивы, если он рассматривает те же вычисления как часть бесконечной последовательности с фиксированным значением $\Delta t/(\Delta z)^2$.

Этот пример, относящийся к любой гиперболической системе, аппроксимация которой выбрана так же, как в (9.19), показывает, что программисту нужна оценка реальной погрешности, а не только уверенность в том, что погрешность будет стремиться к нулю в конце бесконечной последовательности вычислений на идеальной вычислительной машине. К сожалению, реальную (или хотя бы хорошую) оценку погрешности обычно трудно получить. Однако проведенное выше рассуждение об устойчивости вносит некую ясность в вопрос о погрешности. Это рассуждение для разностной системы I (уравнения (9.18)) показывает, что если программист по какой-либо причине не удовлетворен получаемой точностью и решает изменить сетку, уменьшая Δz в два раза, то он должен уменьшить Δt не в два, а в четыре раза, иначе он может увеличить погрешность вместо того, чтобы уменьшить ее.

Далее, практический критерий устойчивости можно получить следующим образом: члены ряда Фурье, в виде которого ищется численное решение, растут как $e^{\gamma t}$, причем для случая I максимальное значение γ приближенно равно

$$\left[\sqrt{1 + \alpha_0^2} - 1 \right] / \Delta t \quad \text{для} \quad k \Delta z \approx \frac{\pi}{2},$$

¹⁾ Норма G равна корню квадратному из нормы G^*G , который является также наибольшим собственным значением матрицы G^*G , так как последняя нормальна.

т. е. достигается для коротковолновых компонент, в то время как максимальное значение γ для истинного решения уравнения (9.17) не может превышать $v\sigma f$. Так как μ_0 лишь немного меньше 1, а величина f вообще незначительно превышает 1, то в качестве грубого критерия мы получим

$$[V\overline{1 + \alpha_0^2} - 1]/\Delta t < v\sigma,$$

или

$$v\Delta t / [\sigma (\Delta z)^2] < 1$$

(множитель $1/2$ опущен, так как в грубом критерии он несуществен). Для практических целей строгие ограничения были бы неудобны, и поэтому мы ищем более удовлетворительные конечноразностные приближения в сферических гармониках. (Аналогичный практический критерий устойчивости для задачи о совместном распространении звука и тепла приводится в следующей главе; там также будет видна необходимость рассматривать конечные значения Δt и Δz , с которыми производятся вычисления. Третий пример будет дан в § 6 гл. 11; он относится к задаче о колебаниях продольно напряженного стержня.)

§ 9.8. Слой; система разностных уравнений II (схема Фридрихса)

Фридрихс [1954] исследовал класс симметричных линейных гиперболических систем, частным случаем которых являются уравнения в сферических гармониках, и указал общий метод получения устойчивой конечноразностной схемы. Для уравнений в сферических гармониках его схема получается из системы I путем замены u_j^n на $(u_{j-1}^n + u_{j+1}^n)/2$:

$$\begin{aligned} \frac{2u_j^{n+1} - u_{j+1}^n - u_{j-1}^n}{2v\Delta t} + M \frac{u_{j+1}^n - u_{j-1}^n}{2\Delta z} + \frac{\sigma}{2} (u_{j+1}^n + u_{j-1}^n) = \\ = \sigma (1 + f) \frac{1}{2} S (u_{j+1}^n + u_{j-1}^n). \end{aligned} \quad (9.21)$$

Это трехточечная формула, содержащая только одну точку сетки, соответствующую моменту t^{n+1} , и две точки, соответствующие моменту времени t^n .

Для этой системы матрица перехода равна

$$G(\Delta t, k) = (I + \varepsilon F) \cos k \Delta z - i\alpha_0 M \sin k \Delta z,$$

где, как и ранее, мы положили

$$v\sigma\Delta t = \varepsilon, \quad (1 + f)S - I = F, \quad \alpha_0 = \frac{v\Delta t}{\Delta z}.$$

Как было показано в § 3.9 и 4.11, устойчивость не теряется при добавлении такого члена как $\varepsilon F \cos k\Delta z = O(\Delta t)$. Оставшаяся часть G , в точности равная $G(0, k)$, является нормальной, так что условие фон Неймана необходимо и достаточно для устойчивости. Кроме того, модули собственных значений $G(0, k)$ равны $(\cos^2 k\Delta z + \alpha_0^2 \mu^2 \sin^2 k\Delta z)^{1/2}$, где μ — собственное значение матрицы M . Если μ_0 снова обозначает максимум модуля собственных значений матрицы M , то это означает, что условие $\alpha_0 \mu_0 \leq 1$ является необходимым и достаточным для устойчивости; иначе говоря, *схема Фридрихса в случае системы II для слоя устойчива тогда и только тогда, когда*

$$\mu_0 (v\Delta t/\Delta z) \leq 1. \quad (9.22)$$

Конечно, этот вывод согласуется и с более общими выводами Фридрихса. Ниже будет показано, что μ_0 немного меньше единицы, и поэтому приведенное выше ограничение будет незначительно слабее условия $v\Delta t \leq \Delta z$.

§ 9.9. Неявные схемы

Для уравнений в сферических гармониках можно построить множество неявных схем. Простейшей из них является четырехточечная схема, использующая две соседние пространственные точки в момент t^n и те же пространственные точки в момент t^{n+1} . Центрируя разности надлежащим образом, мы можем свести погрешность аппроксимации к $O((\Delta t)^2) + O((\Delta z)^2)$, в то время как для систем I и II погрешность аппроксимации составляет $O(\Delta t) + O((\Delta z)^2)$. Получающаяся система является безусловно устойчивой. Другие безусловно устойчивые неявные схемы можно построить по аналогии с соответствующими схемами для уравнения теплопроводности (см. табл. 8.1). Все эти схемы обладают тем недостатком, что они являются неявными по отношению к переменному z и индексу l , который обозначает компоненты u ($l = 0, 1, 2, \dots, L$). Поэтому система уравнений, которую нужно решать на каждом шагу по времени, довольно сложна, и по этой причине неявная схема, насколько известно авторам, не исследовалась. Оказывается, что только система II является подходящей для метода сферических гармоник, но даже и эта система представляет собой худшую из всех схем, применимых для других аппроксимаций уравнения переноса, которые будут рассмотрены далее в этой главе.

§ 9.10. Метод Вика — Чандрасекхара для слоя

В случае симметрии слоя и изотропного рассеяния уравнение переноса принимает вид (9.6), а именно

$$\left(\frac{1}{v} \frac{\partial}{\partial t} + \mu \frac{\partial}{\partial z} + \sigma \right) \psi(z, \mu, t) = \frac{\sigma(1+f)}{2} \int_{-1}^1 \psi(z, \mu', t) d\mu'.$$

Идея метода Вика — Чандрасекхара¹⁾ заключается в том, что выбирают определенную последовательность значений μ , например $\mu_0, \mu_1, \dots, \mu_L$, и строят уравнения для $L+1$ функций $\psi(z, \mu_j, t)$, оценивая для этих $L+1$ значений μ' интеграл в правой части уравнения переноса по возможности более точно.

Как известно, наиболее точной квадратурной формулой по $L+1$ точкам является в некотором смысле формула Гаусса²⁾; именно, для любой функции $F(x)$

$$\frac{1}{2} \int_{-1}^1 F(x) dx \approx \sum_{j=0}^L C_j F(\mu_j),$$

где $\mu_0, \mu_1, \dots, \mu_L$ — нули полинома Лежандра $P_{L+1}(\mu)$, а коэффициенты C_j определяются так:

$$C_j = \left[\sum_{k=0}^L (2k+1) (P_k(\mu_j))^2 \right]^{-1}.$$

В этих обозначениях правая часть уравнения переноса принимает вид

$$\sigma(1+f) \sum_{j=0}^L C_j \psi(z, \mu_j, t).$$

Левую часть уравнения переноса для каждого из $L+1$ значений μ оставим неизменной. Полученную систему $L+1$ дифференциальных уравнений в частных производных лучше всего представить в матрично-векторной форме, как это было сделано в методе сферических гармоник. Пусть $\tilde{\mathbf{u}} = \tilde{\mathbf{u}}(z, t)$ есть $(L+1)$ -мерный вектор

$$\tilde{\mathbf{u}} = \tilde{\mathbf{u}}(z, t) = \begin{pmatrix} \psi(z, \mu_0, t) \\ \psi(z, \mu_1, t) \\ \cdot \\ \cdot \\ \psi(z, \mu_L, t) \end{pmatrix}$$

¹⁾ См. работы Вика [1943] и Чандрасекхара [1944b].

²⁾ По поводу вывода этой формулы и других возможных представлений для коэффициентов см. Хилдебранд [1956].

Тогда уравнения примут вид

$$\left(\frac{1}{v} \frac{\partial}{\partial t} + \tilde{M} \frac{\partial}{\partial z} + \sigma \right) \tilde{u} = \sigma (1 + f) \tilde{S} \tilde{u}, \quad (9.23)$$

где $\tilde{M}$ — диагональная матрица

$$\tilde{M} = \begin{bmatrix} \mu_0 & & & 0 \\ & \mu_1 & & \\ & & \ddots & \\ 0 & & & \mu_L \end{bmatrix},$$

а $\tilde{S}$ — матрица коэффициентов

$$\tilde{S} = \begin{bmatrix} C_0 & C_1 & \dots & C_L \\ C_0 & C_1 & \dots & C_L \\ \cdot & \cdot & \cdot & \cdot \\ C_0 & C_1 & \dots & C_L \end{bmatrix}.$$

Мы увидим, что эти уравнения в случае задачи для слоя более подходят для конечноразностных методов, чем соответствующие уравнения в сферических гармониках. Поэтому желательно обобщить их на задачи со сферической симметрией и анизотропным рассеянием. Это можно сделать, если учесть замечание Дж. К. Марка (неопубликованное сообщение), что в случае задачи для слоя с изотропным рассеянием уравнения в сферических гармониках и уравнения Вика — Чандрасекхара эквивалентны. Последние можно получить из первых, применив к рассматриваемым матрицам некоторое преобразование подобия; если это преобразование подобия применить к более общим уравнениям в сферических гармониках, то мы получим желаемый результат.

Методы такого общего типа, включая описанный ниже S_n -метод Карлсона, использующие диагональную матрицу $\tilde{M}$, называются дискретно-ординатными методами. В серии статей Келлер и Вендрофф [1960] при разумных предположениях доказали сходимость этих методов при $L \rightarrow \infty$.

§ 9.11. Эквивалентность двух методов

Сначала установим (дадим только намек на доказательство) некоторые замечательные свойства матрицы M , входящей в уравнения в сферических гармониках (§ 9.5).

1. Собственные значения матрицы M являются нулями полинома Лежандра $P_{L+1}(\lambda)$. Чтобы убедиться в этом, напомним

$$\mathcal{P}_{L+1}(\lambda) = \frac{(2L+2)!}{2^{L+1} ((L+1)!)^2} \cdot \det(\lambda I - M).$$

Если детерминант разложить по элементам последней строки, то получатся два члена, соответствующие двум отличным от нуля элементам этой строки. Их можно выразить через $\mathcal{P}_L(\lambda)$ и $\mathcal{P}_{L-1}(\lambda)$, которые содержат главные миноры этого детерминанта. Мы получим таким образом рекуррентное соотношение для $\mathcal{P}_L(\lambda)$, равноценное одному из обычных рекуррентных соотношений для полиномов Лежандра. Кроме того, легко видеть, что $\mathcal{P}_1(\lambda) = P_1(\lambda)$ и $\mathcal{P}_2(\lambda) = P_2(\lambda)$. Поэтому $\mathcal{P}_{L+1}(\lambda) = P_{L+1}(\lambda)$, и утверждение доказано.

2. Так как матрица M симметрична, то ясно, что должно существовать ортогональное преобразование, связывающее M и $\tilde{M}$, поскольку матрица $\tilde{M}$ диагональна и имеет те же самые собственные значения, что и M ; можно показать, что

$$U^{-1}MU = \tilde{M}, \text{ где}$$

$$U_{jk} = \sqrt{(2j+1)C_k} P_j(\mu_k).$$

Для этого заметим сначала, что вектор $\mathbf{v}^k$, компоненты которого равны $\sqrt{2j+1} P_j(\mu_k)$ ($j = 0, \dots, L$), удовлетворяет уравнению $(\mu_k I - M)\mathbf{v}^k = 0$ (вследствие рекуррентных соотношений для $P_j(\mu)$). Таким образом получается система собственных векторов матрицы M , и остается только нормировать их¹⁾, чтобы получить матрицу U .

Если метод сферических гармоник и метод Вика — Чандрасекхара эквивалентны, то, вспоминая определение $\psi_j(z, t)$, мы должны ожидать, что

$$\psi(z, \mu_j, t) = \sum_{k=0}^L \psi_k(z, t) \sqrt{2k+1} P_k(\mu_j),$$

или в матричной записи $\tilde{\mathbf{u}} = T\mathbf{u}$, $\mathbf{u} = T^{-1}\tilde{\mathbf{u}}$, где T — матрица, элементы которой имеют вид

$$T_{jk} = \sqrt{2k+1} P_k(\mu_j). \quad (9.24)$$

Если мы применим преобразование T к уравнениям в сферических гармониках, то получим

$$\left(\frac{1}{v} \frac{\partial}{\partial t} + TMT^{-1} \frac{\partial}{\partial z} + \sigma \right) \tilde{\mathbf{u}} = \sigma(1+f) TST^{-1} \tilde{\mathbf{u}},$$

и эквивалентность будет установлена, если доказать, что

$$TMT^{-1} = \tilde{M}, \quad TST^{-1} = \tilde{S}.$$

¹⁾ Выражение для C_k , данное в предыдущем параграфе, очевидно, является условием нормирования для вектора $\mathbf{v}^k$. Используя тождество Дарбу — Кристоффеля, можно представить этот коэффициент в более привычном виде (этот вопрос рассмотрен у Хилдебранда [1956], стр. 322).

Чтобы доказать эти соотношения, заметим, что матрица T тесно связана с матрицей U ; действительно, $T = V^{-1}U^{-1}$, $T^{-1} = UV$, где V — диагональная матрица, для которой $V_{jk} = \sqrt{C_j} \delta_{jk}$. Поэтому

$$TMT^{-1} = V^{-1}U^{-1}MUV = V^{-1}\tilde{M}V = \tilde{M};$$

в этом результате последний знак равенства обусловливается тем, что матрицы V и $\tilde{M}$ диагональны. Заметим, далее, что

$$(T^{-1})_{jk} = C_k \sqrt{2j+1} P_j(\mu_k)$$

и что при изотропном рассеянии $p(\cos \alpha) \equiv 1$, $p_0 = 1$, $p_l = 0$ ($l > 1$), так что единственным ненулевым элементом S является $S_{00} = 1$. Потому

$$(TST^{-1})_{jk} = T_{j0}(T^{-1})_{0k} = C_k = \tilde{S}_{jk}.$$

Этим завершается доказательство того, что для слоя при изотропном рассеянии уравнения Вика — Чандрасекхара и уравнения в сферических гармониках эквивалентны. В терминологии, принятой в теории характеристик (см. § 12.7, а также книгу Куранта [1964]), уравнения Вика — Чандрасекхара будут уравнениями в сферических гармониках, выраженными в нормальной или характеристической форме.

§ 9.12. Граничные условия

В случае слоя, простирающегося от $z = -a$ до $z = a$, граничные условия (9.5) могут быть перенесены в метод Вика — Чандрасекхара без существенных изменений. Пусть μ_j (нули функции $P_{L+1}(\mu)$) расположены так, что $\mu_0 > \mu_1 > \dots > \mu_L$. (Заметим, что μ_0 будет также наибольшим по модулю.) Если $L + 1$ четно, то в качестве граничных условий примем

$$\begin{aligned} \psi(-a, \mu_j, t) &= 0 \quad \text{при } j = 0, 1, \dots, (L-1)/2, \\ \psi(a, \mu_j, t) &= 0 \quad \text{при } j = (L+1)/2, \dots, L. \end{aligned}$$

Если $L + 1$ нечетно, то $\mu_{L/2} = 0$, и граничные условия для $\psi(z, \mu_0, t)$ на первый взгляд кажутся неясными. Однако к уравнению для этой функции не требуется граничное условие, так как в нем отсутствует член с $\partial/\partial z$. Простота этих граничных условий является преимуществом метода Вика — Чандрасекхара. Одним из путей получения граничных условий для метода сферических гармоник является преобразование приведенных выше уравнений к уравнениям в сферических гармониках; это преобразование определяется матрицей T .

§ 9.13. Разностные системы I и II

Конечноразностные уравнения системы I и II практически те же самые, что и в сферических гармониках, и могут быть получены формально из уравнений § 9.6 и 9.8, если символы u , M и S снабдить тильдой. Они имеют те же свойства: система I устойчива при любом фиксированном значении $\Delta t/(\Delta z)^2$, но неустойчива при любом фиксированном значении $\Delta t/\Delta z$ и практически в качестве условия устойчивости может быть принято неравенство $v \Delta t/[\sigma(\Delta z)^2] < 1$; система II (схема Фридрихса) устойчива при $v \Delta t \leq \Delta z/\mu_0$. Помимо этого существуют некоторые системы конечноразностных уравнений, которые не имеют аналогов в методе сферических гармоник. Рассмотрим теперь эти системы.

§ 9.14. Система III; пространственные разности вперед и назад ¹⁾

Для выражения $\partial u/\partial z$ применим при $\mu_j > 0$ пространственное разностное отношение назад, а при $\mu_j < 0$ — отношение вперед. (Если $\mu_j = 0$, что может иметь место при четном L , то производная по пространству отсутствует.) Для этого разобьем $\tilde{M}$ на две части: $\tilde{M} = \tilde{M}_1 + \tilde{M}_2$, где

$$\tilde{M}_1 = \begin{bmatrix} \mu_0 & & & & & & 0 \\ & \cdot & & & & & \\ & & \cdot & & & & \\ & & & \cdot & & & \\ & & & & \mu_{l'} & & \\ & & & & 0 & & \\ & & & & & 0 & \\ & & & & & & 0 \\ & & & & & & \cdot \\ & & & & & & \cdot \\ & & & & & & \cdot \\ & & & & & & 0 \\ 0 & & & & & & 0 \end{bmatrix},$$

$$\tilde{M}_2 = \begin{bmatrix} 0 & & & & & & 0 \\ & 0 & & & & & \\ & & \cdot & & & & \\ & & & \cdot & & & \\ & & & & 0 & & \\ & & & & & \mu_{l''} & \\ & & & & & \cdot & \\ & & & & & \cdot & \\ & & & & & \cdot & \\ & & & & & \cdot & \\ & & & & & \cdot & \\ & & & & & \cdot & \\ 0 & & & & & & \mu_L \end{bmatrix}.$$

¹⁾ Ср. с работой Куранта, Изаксона и Риса [1952].

Здесь μ'_i и μ''_i являются последней положительной и первой отрицательной из величин μ_j соответственно. Разностные уравнения принимают вид

$$\frac{\tilde{u}_j^{n+1} - \tilde{u}_j^n}{v \Delta t} + \tilde{M}_1 \frac{\tilde{u}_j^n - \tilde{u}_{j-1}^n}{\Delta z} + \tilde{M}_2 \frac{\tilde{u}_{j+1}^n - \tilde{u}_j^n}{\Delta z} + \sigma \tilde{u}_j^n = (1 + f) \sigma \tilde{S} \tilde{u}_j^n. \quad (9.25)$$

Эти уравнения легко исследовать на устойчивость применявшимися ранее методами. Для этой системы G представляет собой сумму диагональной матрицы и $O(\Delta t)$. Следовательно, условием фон Неймана для системы (9.25) является неравенство

$$\mu_0 (v \Delta t / \Delta z) \leq 1,$$

которое будет также и достаточным условием устойчивости.

Погрешность аппроксимации для этой системы составляет $O(\Delta t) + O(\Delta z)$, т. е. формально данная система менее точна, чем система Фридрихса. Однако практически она не менее точна, и, вероятно, несколько более точна по следующей причине: в системе Фридрихса при Δt , лежащем вблизи максимально допустимого значения $\Delta z / \mu_0 v$, первый из членов погрешности $O(\Delta t) + O[(\Delta z)^2]$ значительно больше второго, тогда как два члена погрешности системы III сравнимы и имеют разные знаки, в результате чего они частично уничтожаются. Подходящий выбор пространственных разностей вперед и назад в зависимости от знака μ в известной мере позволяет следовать характеристикам ближе, чем в схеме Фридрихса. Обе схемы имеют одинаковые условия устойчивости.

§ 9.15. Система IV (неявная)

Исходя из опытов с уравнением теплопроводности, можно ожидать, что устойчивость системы (9.25) значительно возрастет, если мы используем разности по времени назад (или, что почти то же самое, заменим n на $n + 1$ в индексах пространственных разностей). Для того чтобы следовать характеристикам возможно ближе, мы также поменяем роли¹⁾ $\tilde{M}_1$ и $\tilde{M}_2$. В результате получим систему

$$\frac{\tilde{u}_j^{n+1} - \tilde{u}_j^n}{v \Delta t} + \tilde{M}_1 \frac{\tilde{u}_{j+1}^{n+1} - \tilde{u}_j^{n+1}}{\Delta z} + \tilde{M}_2 \frac{\tilde{u}_j^{n+1} - \tilde{u}_{j-1}^{n+1}}{\Delta z} + \sigma \tilde{u}_j^n = (1 + f) \sigma \tilde{S} \tilde{u}_j^n. \quad (9.26)$$

¹⁾ Если $\tilde{M}_1$ и $\tilde{M}_2$ не менять ролями, то полученные уравнения окажутся безусловно устойчивыми, но совершенно не будут следовать характеристикам.

Матрица перехода для этой системы может быть легко вычислена. Вопреки ожиданию, данная система не является безусловно устойчивой. В самом деле, *условием устойчивости является неравенство*

$$\mu^* (v \Delta t / \Delta z) \geq 1,$$

где μ^* — наименьший положительный нуль $P_{L+1}(\mu)$. Отметим необычное свойство этой системы: именно условие устойчивости ограничивает Δt снизу, а не сверху. Вследствие малости μ^* эта система практически не используется.

§ 9.16. Система V (схема Карлсона)

В связи со своим S_n -методом, который будет кратко описан ниже и напоминает в некоторых отношениях метод Вика — Чандрасекхара, Б. Карлсон [1953] получил безусловно устойчивую конечноразностную систему; эта схема сочетает лучшие свойства системы III (уравнение (9.25)) и системы IV (уравнение (9.26)). Мы изложим несколько упрощенный вариант этой схемы в том виде, в каком она применяется в методе Вика — Чандрасекхара. Идея Карлсона заключается в рассмотрении компонент $\tilde{u}$, соответствующих схеме (9.25) для малых μ и схеме (9.26) для больших μ . Предположим, что величина $v\Delta t/\Delta z$ является постоянной, и разобьем диагональную матрицу $\tilde{M}$ на четыре слагаемых: $\tilde{M} = \tilde{M}_3 + \tilde{M}_4 + \tilde{M}_5 + \tilde{M}_6$, где

$\tilde{M}_3$ содержит те диагональные элементы $\tilde{M}_1$, для которых $\mu > \Delta z/(v \Delta t)$, $\tilde{M}_4$ содержит те диагональные элементы $\tilde{M}_1$, для которых $\mu \leq \Delta z/(v \Delta t)$, $\tilde{M}_5$ содержит те диагональные элементы $\tilde{M}_2$, для которых $\mu \geq -\Delta z/(v \Delta t)$, $\tilde{M}_6$ содержит те диагональные элементы $\tilde{M}_2$, для которых $\mu < -\Delta z/(v \Delta t)$; все прочие элементы являются нулями в каждом случае.

Разностные уравнения теперь примут следующий вид:

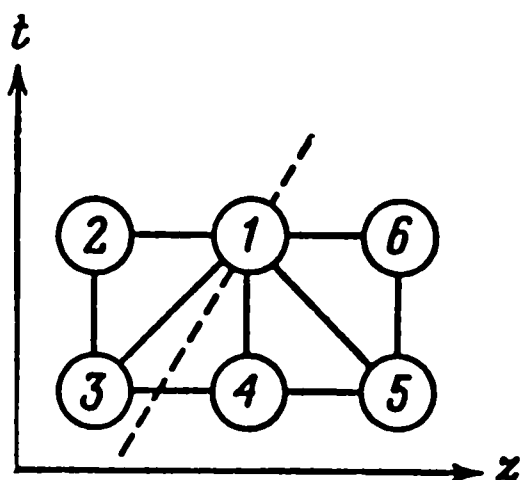
$$\begin{aligned} \frac{\tilde{u}_j^{n+1} - \tilde{u}_j^n}{v \Delta t} + \tilde{M}_3 \frac{\tilde{u}_{j+1}^{n+1} - \tilde{u}_j^{n+1}}{\Delta z} + \tilde{M}_4 \frac{\tilde{u}_j^n - \tilde{u}_{j-1}^n}{\Delta z} + \tilde{M}_5 \frac{\tilde{u}_{j+1}^n - \tilde{u}_j^n}{\Delta z} + \\ + \tilde{M}_6 \frac{\tilde{u}_j^{n+1} - \tilde{u}_{j-1}^{n+1}}{\Delta z} + \sigma \tilde{u}_j^n = (1 + f) \sigma \tilde{S} \tilde{u}_j^n. \end{aligned} \quad (9.27)$$

На рис. 9.1 изображена эта схема и дана ее связь с характеристиками. Показаны шесть соседних точек сетки в плоскости z, t ; верхний ряд соответствует времени t^{n+1} , а нижний ряд — времени t^n . Характеристиками являются $L + 1$ прямых; наклон l -й характеристики равен $1/(\mu_l v)$. В зависимости от величины этого наклона характеристика, проходящая через точку 1, может лежать внутри одного из четырех треугольников, изображенных на рисунке. Для l -й компоненты вектора $\tilde{u}$ дифференциальный оператор в уравнении (9.23) имеет вид $(1/v) \partial/\partial t +$

$+ \mu_l \partial / \partial z$, т. е. является оператором дифференцирования вдоль l -й характеристики. В схеме Карлсона производная выражается через значения функции в трех точках: в точке 1 и в одной из пар точек (2, 3), (3, 4), (4, 5) или (5, 6) соответственно тому, в каком из четырех треугольников лежит характеристика. Если одна из характеристик проходит через точку 3, то соответствующий конечноразностный оператор автоматически сводится к двухточечной формуле, использующей точки 1 и 3, и имеет вид

$$(\tilde{u}_j^{n+1} - \tilde{u}_{j-1}^n) / v \Delta t.$$

В этом случае производная вдоль характеристики заменяется на разностное отношение вдоль нее. Это, конечно, лучше чем



Р и с. 9.1. Схема расположения шести точек сетки в плоскости (z, t) , иллюстрирующая метод Карлсона. Штриховой линией показана типичная характеристика.

менее прямой способ, в котором применяются частные производные и используется третья точка. Подобное замечание относится, очевидно, и к характеристикам, проходящим через точки 4 или 5. Для характеристики, не проходящей ни через одну из этих точек, необходимо использовать (по меньшей мере) три точки для выражения дифференциального оператора. Достоинство схемы Карлсона заключается в том, что для каждой компоненты u выбираются три точки, лежащие возможно ближе к соответствующей характеристике.

Как и в предыдущих схемах, матрица перехода G и условие устойчивости получаются без особого труда. В результате оказывается, что *схема Карлсона (9.27) устойчива при любом значении величины $v \Delta t / \Delta z$.*

Хотя схема Карлсона представляет собой неявную схему, можно легко получить численное решение следующим образом¹⁾. При $\mu_l < 0$ некоторое значение (а именно нуль) $\psi(z, \mu_l, t^{n+1})$ задано граничным условием на правой стороне слоя при $z = a$. Разностные уравнения можно решать, уменьшая j

¹⁾ Для увеличения точности часто u_j^n в недифференциальных членах заменяют на $1/2(u_j^n + u_j^{n+1})$ (это не влияет на устойчивость). Тогда неявные уравнения не удастся решить так просто, и в общем случае возникает необходимость в итерационном процессе.

(уменьшая z) до тех пор, пока не будет достигнута левая граница. Аналогично, при $\mu_l > 0$ граничное значение задано при $z = -a$, и уравнения можно решать в порядке возрастания j . При $\mu_l = 0$ (что имеет место при четном L) уравнения становятся явными.

Применение схемы Карлсона к более общим квазилинейным гиперболическим системам было рассмотрено Келлером [1956]. Так как в случае переменных коэффициентов наклон характеристик может меняться от одной точки к другой, то в процессе интегрирования необходимо видоизменять схему в зависимости от положения характеристики (четыре положения). Тем не менее возможно, следуя Келлеру, решать разностные уравнения при помощи несложного алгоритма. Келлер доказал также устойчивость и сходимость при разумных предположениях. Келлер и Вендрофф [1957] провели подробное исследование этой схемы для широкого класса дискретно-ординатных методов; они установили безусловную устойчивость этой схемы, а также сходимость итерационного процесса, упомянутого в последнем подстрочном примечании.

§ 9.17. Обобщение метода Вика — Чандрасекхара

В § 10 предполагалось, что имеет место изотропное рассеяние и симметрия слоя. Если рассеяние анизотропно, то матрицу $\tilde{S}$ можно заменить матрицей TST^{-1} , где S — общая матрица, данная в § 9.5 для метода сферических гармоник, а T — матрица преобразования, определенная в § 9.11. Таким образом, матрица $\tilde{S}$ берется в виде

$$(\tilde{S})_{jk} = C_k \sum_l P_l(\mu_j) p_l P_l(\mu_k)$$

и является обобщением матрицы S , рассмотренной в § 9.10. Множители p_l являются коэффициентами в формуле рассеяния (9.10). Все упомянутые конечноразностные схемы применимы без всякого изменения устойчивости, так как при исследовании устойчивости не было сделано предположений о каком-либо специальном виде матрицы $\tilde{S}$.

Для обобщения метода на задачи со сферической симметрией включим в левую часть уравнения (9.23) член $(1/r)\tilde{N}\tilde{u}$, где $\tilde{N}$ — соответственно преобразованная матрица N , входящая в уравнения сферических гармоник (см. § 9.5). Таким образом,

$$\begin{aligned} (\tilde{N})_{jk} &= (TNT^{-1})_{jk} = \\ &= C_k \sum_{l=0}^{L-1} (l+1) [(l+2) P_l(\mu_j) P_{l+1}(\mu_k) - l P_l(\mu_k) P_{l+1}(\mu_j)]. \end{aligned}$$

При добавлении члена $(1/r)\tilde{N}\tilde{u}$ — наш критерий устойчивости не будет строго справедливым, так как теперь коэффициенты переменны. Но если следовать обычному эвристическому подходу, согласно которому устойчивость является локальным явлением (так что при медленном изменении коэффициентов их можно считать постоянными в малой окрестности), то следует ожидать, что добавление таких членов не повлияет на устойчивость, так как при r_0 , равном константе, добавление члена $(1/r_0)\tilde{N}\tilde{u}$ оставляет условие устойчивости неизменным для всех рассмотренных выше конечноразностных схем. Насколько нам известно, этот метод не применялся для нестационарных задач, и мы предпочитаем воздержаться от суждения о том, что может происходить вблизи центра в связи с имеющейся там особенностью. Однако практика работы с S_n -уравнениями, которые подобны рассматриваемым уравнениям, наводит на мысль о том, что здесь, вероятно, не возникают каких-либо трудностей. Снова может быть применена схема Карлсона с интегрированием от границы к центру при $\mu < 0$, а затем от центра назад к границе при $\mu > 0$.

§ 9.18 S_n -метод Карлсона [1953]

Этот метод похож на метод Вика — Чандрасекхара, однако не эквивалентен ему. Он отличается тем, что производная по μ (которая фигурирует в задачах со сферической симметрией) аппроксимируется более прямым способом; кроме того, существует определенное отличие в центрировании по μ , приводящее к тому, что некоторые матрицы, диагональные в методе Вика — Чандрасекхара, здесь оказываются только треугольными.

Разделим отрезок $-1 \leq \mu \leq 1$ на n подинтервалов (не обязательно равных):

$$-1 = \mu_0 < \mu_1 < \dots < \mu_n = +1,$$

причем величины $\mu_0, \dots, \mu_n$ в общем случае не совпадают с величинами $\mu_0, \dots, \mu_L$, введенными ранее; далее будут написаны $n+1$ уравнений в частных производных для $n+1$ функций $N_0(r, t) = \psi(r, \mu_0, t)$, $N_1(r, t) = \psi(r, \mu_1, t), \dots, N_n(r, t) = \psi(r, \mu_n, t)$. Каждое из этих уравнений, за исключением первого, центрируется относительно средней точки одного из подинтервалов.

Рассмотрим случай сферической симметричной системы с изотропным рассеянием. Уравнение переноса записывается (ср. с уравнением (9.7)) в следующем виде:

$$\left(\frac{1}{v} \frac{\partial}{\partial t} + \mu \frac{\partial}{\partial r} + \frac{1-\mu^2}{r} \frac{\partial}{\partial \mu} + \sigma \right) \psi(r, \mu, t) = \sigma \frac{1+f}{2} \int_{-1}^1 \psi(r, \mu', t) d\mu'.$$

Пусть интеграл аппроксимируется следующим образом:

$$\frac{1}{2} \int_{-1}^1 \psi(r, \mu', t) d\mu' \approx \sum_{i=0}^n A_i N_i(r, t),$$

где коэффициенты A_i соответствуют принятой формуле приближенного интегрирования, например формуле трапеций.

При $\mu = \pm 1$ производная по μ не входит в уравнение переноса, и в качестве уравнения для $N_0(r, t)$, которое соответствует $\mu = -1$, мы можем взять просто

$$\left(\frac{1}{v} \frac{\partial}{\partial t} - \frac{\partial}{\partial r} + \sigma \right) N_0(r, t) = \sum_{i=1}^n A_i N_i(r, t). \quad (9.28)$$

Для уравнения, которое центрировано относительно средней точки подинтервала (μ_i, μ_{i+1}) , производная по времени и член $\sigma\psi$ заменяются простыми средними, а другие члены аппроксимируются более сложными выражениями:

$$\begin{aligned} \frac{1}{2v} \left(\frac{\partial N_i}{\partial t} + \frac{\partial N_{i+1}}{\partial t} \right) + \frac{1}{2} \left(\mu_i^+ \frac{\partial N_i}{\partial r} + \mu_{i+1}^- \frac{\partial N_{i+1}}{\partial r} \right) + \\ + \frac{1 - (\overline{\mu^2})_{i+1/2}}{r} \frac{N_{i+1} - N_i}{\mu_{i+1} - \mu_i} + \frac{\sigma}{2} (N_i + N_{i+1}) = \\ = \sigma (1 + f) \sum_{p=0}^n A_p N_p; \end{aligned} \quad (9.29)$$

здесь μ_i^+ и μ_{i+1}^- — должным образом выбранные точки подинтервала (μ_i, μ_{i+1}) вблизи его левого и правого концов соответственно, а $(\overline{\mu^2})_{i+1/2}$ — надлежаще выбранное усреднение величины μ^2 в этом подинтервале.

Для S_n -метода, кроме того, характерен специальный выбор значений μ_i^+ , μ_{i+1}^- и $(\overline{\mu^2})_{i+1/2}$, а именно

$$\begin{aligned} \mu_i^+ &= \frac{1}{3} (2\mu_i + \mu_{i+1}), \quad \mu_{i+1}^- = \frac{1}{3} (\mu_i + 2\mu_{i+1}), \\ (\overline{\mu^2})_{i+1/2} &= \frac{1}{3} (\mu_i^2 + \mu_i \mu_{i+1} + \mu_{i+1}^2). \end{aligned}$$

Этот выбор обосновывается Карлсоном отчасти эвристически и базируется на том обстоятельстве, что при таком выборе и при предположении линейной зависимости ψ от μ в пределах подинтервала мы получаем в уравнении (9.29) члены, совпадающие со средними по μ в этом подинтервале величинами

соответствующих членов уравнения переноса. Келлер и Вендрофф [1957] выбирали эти значения несколько иначе, полагая $\mu_i^+ = \mu_{i+1}^-$, так что система (9.29) при этом записывается в нормальном или характеристическом виде; это дает возможность рассмотреть вопрос об устойчивости теоретически.

Для функции $N_n(r, t)$, соответствующей $\mu = 1$, можно написать уравнение, аналогичное уравнению (9.28). В этом уравнении нет необходимости, однако оно может быть использовано для проверки точности результатов.

С точностью до центрирования недифференциальных величин по времени конечноразностная схема по r, t , применяемая для решения этих уравнений, совпадает с той, которую ранее в этой главе (§ 9.16 и 9.17) мы называли схемой Карлсона. Расчет помогает определить, что дает замена таких величин средними: например, в разностных уравнениях, которые позволяют продолжить решение от момента времени t^v до момента t^{v+1} , слагаемое $1/2 \sigma N_i(r, t)$ заменяется на $1/4 \sigma [N_i(r, t^v) + N_i(r, t^{v+1})]$. Нельзя формально оправдать такую замену, так как погрешность аппроксимации для схемы Карлсона в любом случае будет равна $O(\Delta t) + O(\Delta r)$, однако так как схема Карлсона в определенном смысле следует характеристикам как можно ближе, то происходит частичное погашение членов $O(\Delta t)$ и $O(\Delta r)$, обусловленных производными. Поэтому кажется полезным задавать другие члены с точностью до $O(\Delta t)^2$.

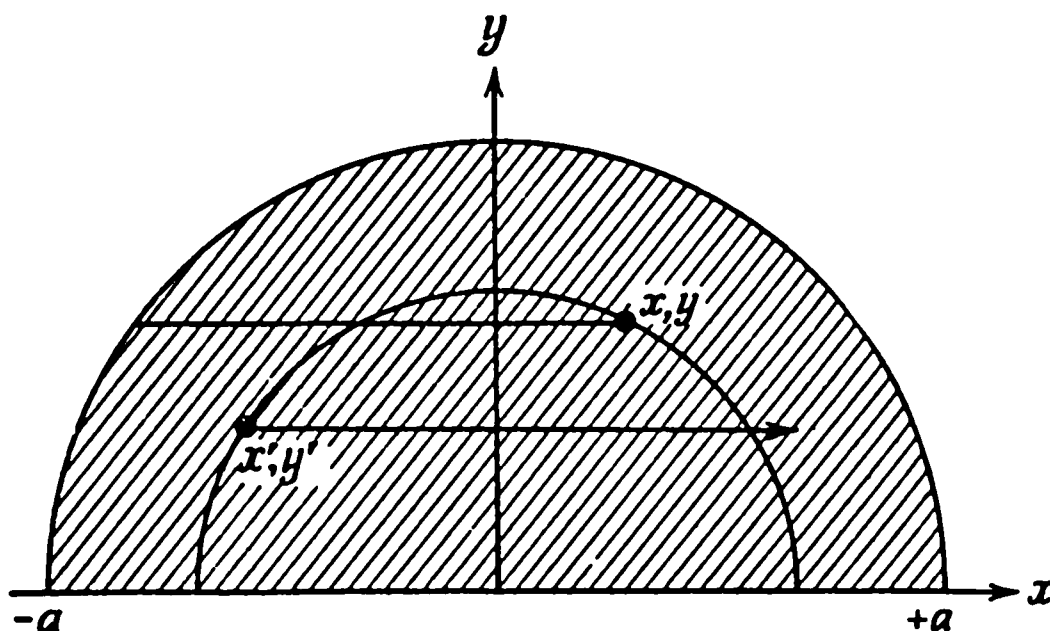
Центрирование по времени можно без труда осуществить во всех членах левой части уравнения (9.29). Для центрирования же сумм правой части необходимо знать все неизвестные функции $N_0(r, t^{v+1}), \dots, N_n(r, t^{v+1})$ в каждом уравнении. Это препятствует нахождению решения путем простого интегрирования от границы к центру при $\mu < 0$ и затем от центра снова к границе при $\mu > 0$. Поэтому используется следующий итерационный процесс: сначала для получения предварительных значений неизвестных величин на момент времени t^{v+1} решаем уравнения с нецентрированной правой частью; затем подставляем результат в (центрированный) правый член и уточняем значения неизвестных и т. д. Подобный итерационный процесс будет описан более подробно в § 9.19 для метода прямого интегрирования при решении уравнения переноса. Как было отмечено выше, Келлер и Вендрофф [1957] доказали сходимость этого итерационного процесса для своего варианта S_n -метода.

Карлсон рассмотрел устойчивость своей схемы в применении к S_n -уравнениям. Из его заключений и результатов большого количества расчетов, проведенных по его методу, следует, что схема является безусловно устойчивой, так же как и для метода Вика — Чандрасекхара. При выборе коэффициентов в уравнении (9.29) по Карлсону аналог матриц M и $\tilde{M}$, кото-

рые рассматривались в описанных выше методах, не будет нормальной матрицей, так что исследование устойчивости является довольно сложным. На практике устойчивость этого метода не вызывает сомнений.

§ 9.19. Метод прямого интегрирования

В ряде конечноразностных схем для интегрирования как стационарных, так и зависящих от времени задач переноса со сферической симметрией вместо r и θ используются переменные $x = r \cos \theta$, $y = r \sin \theta$, предложенные одному из авторов



Р и с. 9.2. Область задания уравнения переноса вида (9.30) в плоскости x, y .

фон Нейманом в 1948 г. В этом параграфе детально описывается один таковой метод для иллюстрации некоторых основных положений.

Если в момент времени t нейтрон движется по прямой, то переменная y является кратчайшим расстоянием от начала координат до этой прямой, а x — расстоянием от основания опущенного на эту прямую из начала координат перпендикуляра до нейтрона, взятым со знаком, учитывающим направление движения нейтрона (см. рис. 9.2). Используя эти переменные и полагая $\Psi(x, y, t) = \psi(r, \cos \theta, t)$, получаем уравнение переноса в виде

$$\begin{aligned} \left(\frac{1}{v} \frac{\partial}{\partial t} + \frac{\partial}{\partial x} + \sigma \right) \Psi(x, y, t) = \\ = \sigma \frac{1 + f}{4\pi r} \int_{-r}^r dx' \int_0^{2\pi} p(\cos \alpha) \Psi(x', \sqrt{r^2 - x'^2}, t) d\varphi \\ \text{при } x^2 + y^2 \leq a^2, \quad y \geq 0, \quad t \geq 0, \end{aligned} \quad (9.30)$$

где r теперь обозначает $\sqrt{x^2 + y^2}$, а α определяется из равенства

$$\cos \alpha = \frac{xx' + y \sqrt{r^2 - x'^2} \cos \varphi}{r^2}.$$

Граничные условия можно записать следующим образом:

$$\Psi(x, y, t) \text{ при } x^2 + y^2 = a^2, \quad y \geq 0, \quad x \leq 0. \quad (9.31)$$

Уравнение переноса справедливо в области заштрихованного полукруга, изображенного на рис. 9.2, а путь интегрирования в уравнении (9.30) представляет собой дугу полуокружности, проходящую через точку (x, y) . На этом рисунке точка изображает нейтрон, движущийся горизонтально с постоянной скоростью v в правую сторону до момента соударения. В момент соударения точка, изображающая нейтрон, переходит скачком в некоторую другую точку той же самой окружности (например, в точку (x', y')), от которой затем продолжает движение горизонтально вправо со скоростью v . Если нейтрон уходит из системы, то он пересекает правую четверть граничной окружности, и граничное условие при отсутствии потока нейтронов извне состоит в том, что $\Psi = 0$ на левой четверти граничной окружности. Методы, основанные на применении этих переменных, иногда называются методами прямого интегрирования, хотя в действительности они являются не более прямыми, чем, например, S_n -метод.

Иногда в плоскости x, y выбирается прямоугольная сетка точек, а иногда в качестве точек сетки берутся точки на пересечении семейства окружностей и семейства горизонтальных прямых. Во всяком случае, точки располагаются в ряды параллельно оси x , поскольку в уравнение входит производная по координате x и не входит производная по координате y . Следовательно, интегрирование производится вдоль пути, по которому нейтроны двигаются между соударениями.

В методе, который здесь будет описан, эта идея получает дальнейшее развитие: производится интегрирование вдоль траекторий нейтронов в пространстве координат-времени по аналогии с методом характеристик для гидродинамических задач. Выберем трехмерную сетку следующим образом:

$$\begin{aligned} x &= i \Delta x, \quad i = -I, -(I-1), \dots, 0, 1, \dots, I, \\ y &= j \Delta y, \quad j = 0, 1, \dots, J, \\ t &= n \Delta t, \quad n = 0, 1, \dots, \end{aligned}$$

где $v \Delta t = \Delta x$ и $I \Delta x = J \Delta y = a$.

Обозначим $\Psi(i\Delta x, j\Delta y, n\Delta t)$ через Ψ_{ij}^n . Ограничимся рассмотрением случая изотропного рассеяния и обозначим правую часть уравнения (9.30) через $\Phi(r, t)$, так что

$$\Phi(r, t) = \sigma \frac{1+f}{2r} \int_{-r}^r \Psi(x', \sqrt{r^2 - x'^2}, t) dx'. \quad (9.32)$$

Далее, будем обозначать $\Phi(r, t^n)$ через $\Phi^n(r)$. Для табулирования Φ выбирается интервал Δr и $\Phi(p\Delta r, n\Delta t)$ обозначается через Φ_p^n . Удобно также для $\sqrt{(i\Delta x)^2 + (j\Delta y)^2}$ ввести обозначение r_{ij} .

Ввиду соотношения $v\Delta t = \Delta x$ точка $((i+1)\Delta x, j\Delta y, (n+1)\Delta t)$ пространства координат-времени находится на той же траектории (характеристике), что и точка $(i\Delta x, j\Delta y, n\Delta t)$. Аппроксимируем члены с производными следующим образом:

$$\left[\left(\frac{1}{v} \frac{\partial}{\partial t} + \frac{\partial}{\partial x} \right) \Psi \right]_{(i+1/2)\Delta x, j\Delta y, (n+1/2)\Delta t} = \frac{\Psi_{i+1,j}^{n+1} - \Psi_{ij}^n}{\Delta x} + O(\Delta t^2)$$

и, наконец, получим

$$\frac{\Psi_{i+1,j}^{n+1} - \Psi_{ij}^n}{\Delta x} + \sigma \frac{\Psi_{i+1,j}^{n+1} + \Psi_{ij}^n}{2} = \Phi^{n+1/2}(r_{i+1/2,j}), \quad (9.33)$$

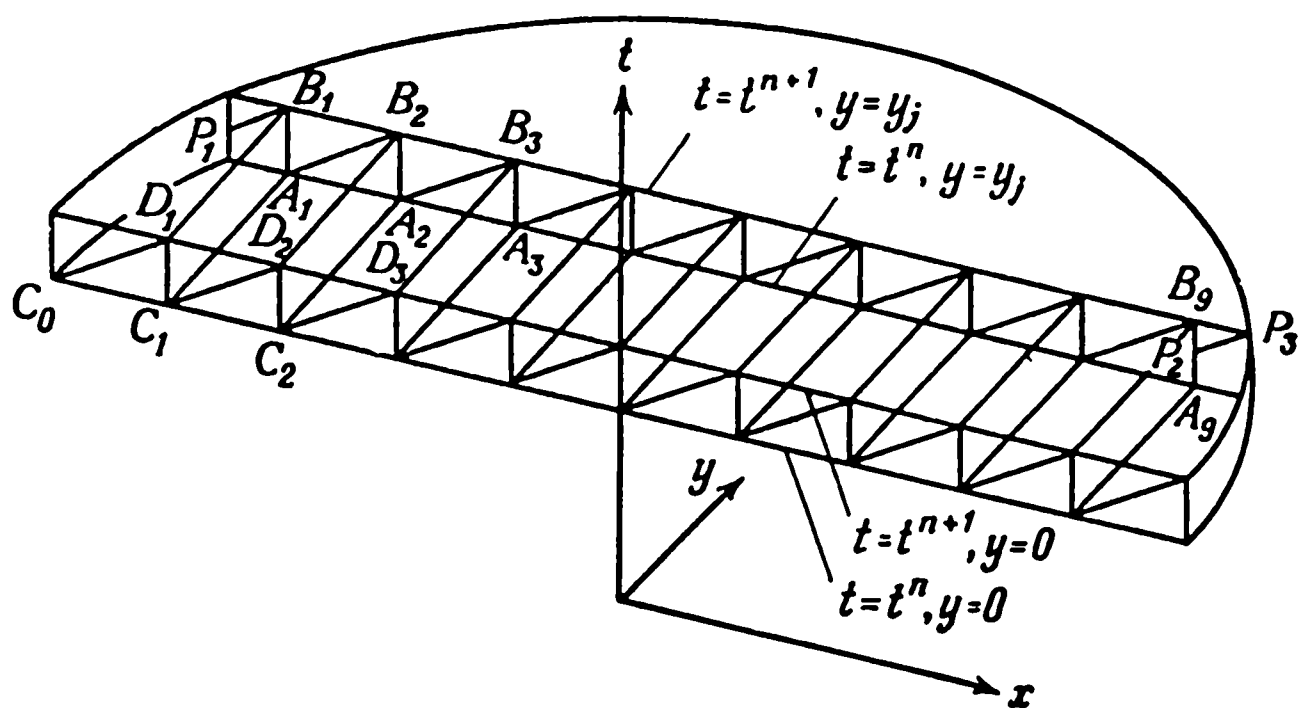
где $\Phi^{n+1/2}(r)$ — приближенное значение $\Phi(r, t)$ в момент времени $(n+1/2)\Delta t$, которое должно быть получено при помощи итерационного процесса, описанного ниже. Если известны значения $\Phi^{n+1/2}$, то из уравнений (9.33) можно найти неизвестные Ψ^{n+1} во всех точках сетки, за исключением точки $(i+1, j)$, обозначающих самую левую точку внутри полукруга на j -й линии (например, точка B_1 на рис. 9.3), в которой величину Φ_{ij}^n нельзя вычислить. Однако согласно граничному условию функция Ψ равна нулю в точке P_1 пространства координат-времени, в которой траектория, проходящая через точку $(i+1, j, n+1)$, входит в систему при радиусе $r = a$. В этом случае мы заменим Δx в уравнениях (9.33) величиной

$$\sqrt{a^2 - (j\Delta y)^2} - |i+1|\Delta x,$$

являющейся расстоянием от точки $(i+1, j)$ до границы вдоль линии $y = j\Delta y$ и $t = \text{const}$, и заменяем Ψ_{ij}^n нулем; тогда уравнения (9.33) дают $\Psi_{i+1,j}^{n+1}$, как и для других точек сетки внутри системы.

На рис. 9.3 показаны характерные особенности трехмерной сетки точек. Отрезки прямых A_1B_2 , A_2B_3 , C_1D_2 и т. д. являются

характеристиками, соединяющими точки сетки для времени t^n с точками для времени t^{n+1} . Правая часть уравнения (9.33) является средним значением величины $\Phi(r, t)$ по такому отрезку и приравнивается среднему от значений $\Phi(r, t)$ на двух концах отрезка. Для отрезка типа A_1B_2 используется значение $\Phi^n(r)$ на другом конце. Для коротких отрезков частного вида P_1B_1 и P_2P_3 уравнение видоизменяется: значение Φ для ближнего конца получается интерполированием по времени между



Р и с. 9.3. Перспективное изображение части трехмерной сетки для уравнения (9.33).

t^n и t^{n+1} . Таким образом, если рассматривать точки A_9 , P_2 и B_9 , то значение $\Phi(P_2)$ получается интерполированием между $\Phi(A_9)$ и $\Phi(B_9)$.

Если значения Ψ_{ij}^{n+1} известны во всех точках сетки, то могут быть найдены Φ_r^{n+1} . Уравнение (9.32) показывает, что Φ_r^{n+1} выражается криволинейным интегралом от Ψ^{n+1} по полуокружности радиуса $\rho \Delta r$ в плоскости x, y . Вообще говоря, эта полуокружность может не пройти ни через одну точку сетки или пройти через небольшое число таких точек, и тогда возникает необходимость в интерполяции. Криволинейный интеграл проще всего берется при интегрировании по x' и использовании формулы трапеции; эта операция легко программируется для машин. При этом в качестве приращения x' используется Δx или $\Delta x/2$, а значения подинтегральной функции в нужных точках переменного x' получаются путем линейной интерполяции по x и y . (Шаг $\Delta x/2$ использовался в пробных вычислениях, описанных ниже.)

Для нахождения $\Phi^{n+1/2}$ применяется следующий итерационный процесс. В начале полного цикла вычислений Ψ_{ij}^n известны и имеется таблица для Φ_r^n . Вместо $\Phi^{n+1/2}(r_{i+1/2, j})$ в пра-

вой части (9.33) берется сначала $\Phi^n(r_{i+1/2, j})$, полученное линейной интерполяцией из таблицы Φ_p^n при $r = r_{i+1/2, j}$. Потом решаются уравнения (9.33) и получаются предварительные значения Ψ_{ij}^{n+1} . С их помощью описанным выше способом получаются предварительные значения для величин Φ_p^{n+1} . Затем выполняется вторая итерация, для которой $\Phi^{n+1/2}$ берется как среднее от Φ^n и Φ^{n+1} . Эта итерация уточняет величины Ψ_{ij}^{n+1} , из которых получаются более точные значения Φ_p^{n+1} .

Опыт показывает, что обычно дальнейшие итерации не имеют смысла, так как применение более мелкой сетки быстрее приводит к цели. Напротив, первая итерация обычно приводит к более точному результату, чем это происходит в результате такой же дополнительной затраты труда для дальнейшего дробления сетки. Вероятно, это происходит потому, что первая итерация уменьшает погрешность аппроксимации от $O(\Delta t)$ до $O(\Delta t^2)$, после чего не происходит дальнейшего изменения порядка величины погрешности.

Этот метод иллюстрируется пробными вычислениями, выполненными на машине УНИВАК. Решается одна из упрощенных астрофизических задач, выбранная в основном для проверки метода. Она описывает движение фотонов, а не движение нейтронов, и поэтому мы обозначим скорость через c , а не через v . В момент времени $t = 0$ в центре большого однородного сферического облака, состоящего из чисто рассеивающего вещества ($f = 0$), происходит короткая вспышка света (как в переменных звездах). При этом предполагается, что рассеяние изотропно и поляризационные эффекты отсутствуют. Требуется найти интенсивность света, испускаемого поверхностью облака, в зависимости от времени.

Было выполнено несколько расчетов для сферы с радиусом, равным удвоенной средней длине свободного пробега, и один расчет для сферы с радиусом, равным учетверенной средней длине свободного пробега. Некоторые результаты представлены графически на рис. 9.4—9.7. Выбрав в качестве единиц длины и времени среднюю длину и среднее время свободного пробега, имеем

случай I: $\sigma = 1, \quad v = c = 1, \quad a = 2;$

случай II: $\sigma = 1, \quad v = c = 1, \quad a = 4.$

В первых двух попытках просчитать случай I ($I = J = 10$, 162 внутренние точки сетки и $I = J = 15$, около 350 внутренних точек сетки) задача рассматривалась просто как задача с начальным условием: при $n = 0$ ($t = 0$) функция Ψ_{ij}^n была равна нулю всюду, кроме центральной точки сетки $i = j = 0$, где

она была произвольно взята равной 0,1. Таким образом, фотоны сначала концентрировались в непосредственной близости от источника, а затем начинали движение в соответствии с уравнением переноса. Этот подход оказался неудачным и был пригодным лишь для грубого качественного исследования из-за разрывного характера распределения в начале движения. Была также предпринята попытка «размазать» начальное распределение по трем или четырем точкам сетки, но и это привело лишь к небольшому улучшению.

Затем схема расчета была видоизменена следующим образом: функция Ψ_{ij}^n приравнивалась нулю при $n = 0$, а член, соответствующий источнику, включался в уравнение переноса в интервале времени $0 \leq t \leq a/c$, в течение которого первоначальный импульс проходил через сферу. В момент времени t источник занимал положение $r = ct$ и представлял фотоны, излученные после первого рассеяния. Иначе говоря, первое рассеяние было рассмотрено аналитически, а последующие — численным способом. Первоначальный импульс расходился в виде расширяющейся сферической оболочки с количеством фотонов на единицу поверхности оболочки, пропорциональным $e^{-\sigma r}/r^2$. Чтобы получить такую пропорциональность для оболочечного источника, необходимо добавить к Φ_r^{n+1} величину $Ae^{-\sigma r} \Delta r/p^2$ после цикла, для которого $n + 1 = p$. Перед этим первым циклом величина Φ_0^0 приближенно представляла фотоны, введенные в систему оболочечным источником в течение интервала времени $(0, \Delta t/2)$, а другие Φ_r^0 были равны нулю. Постоянная A выбиралась пропорциональной I^3 , так чтобы при пересчете задачи с более мелкой сеткой полное количество фотонов, введенных в систему в течение интервала $(0, a/c)$, оставалось постоянным.

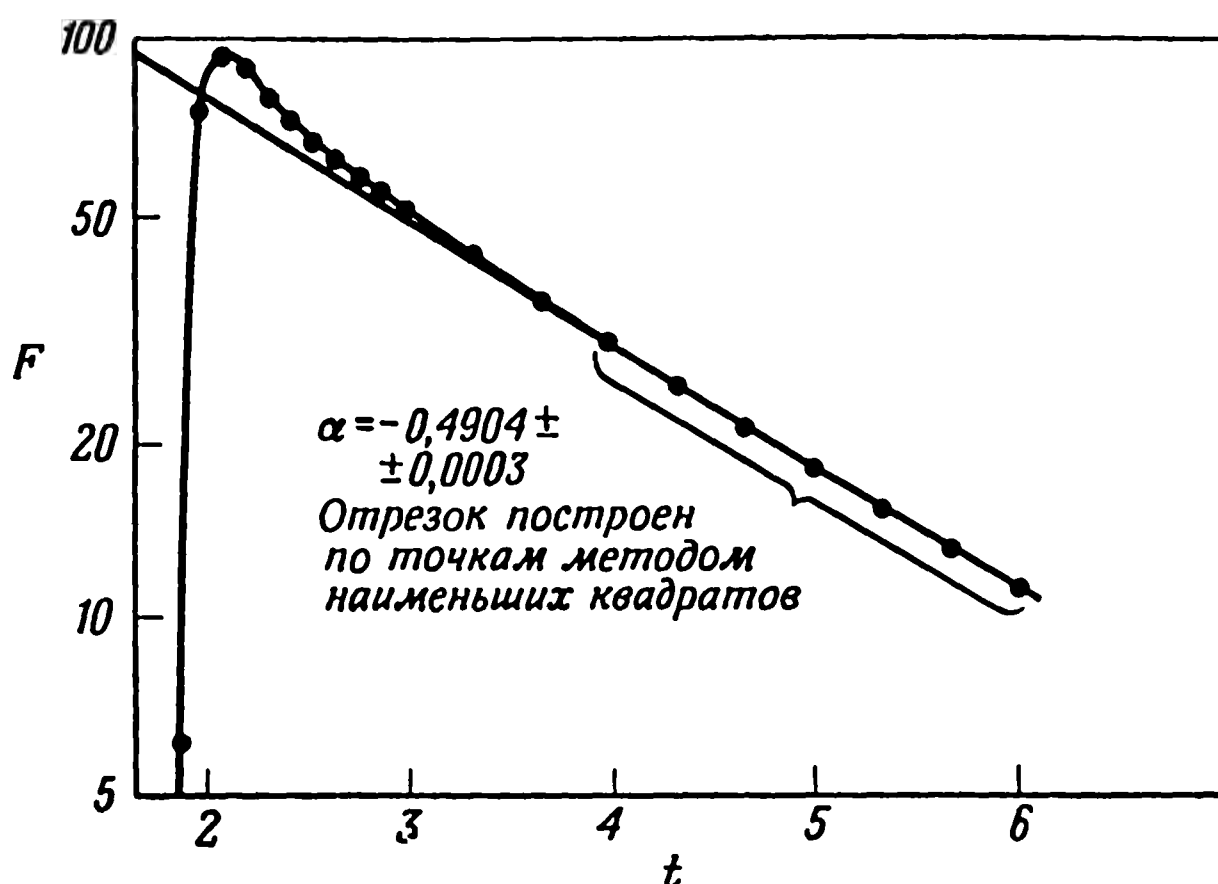
При такой модификации функции Φ и Ψ , которые нужно найти в результате вычислений, остаются разрывными в интервале времени $(0, a/c)$, но они ограничены и для их выражения нет необходимости использовать δ -функцию Дирака.

Световой поток F при $r = a$ получается путем вычисления по формуле трапеций интеграла

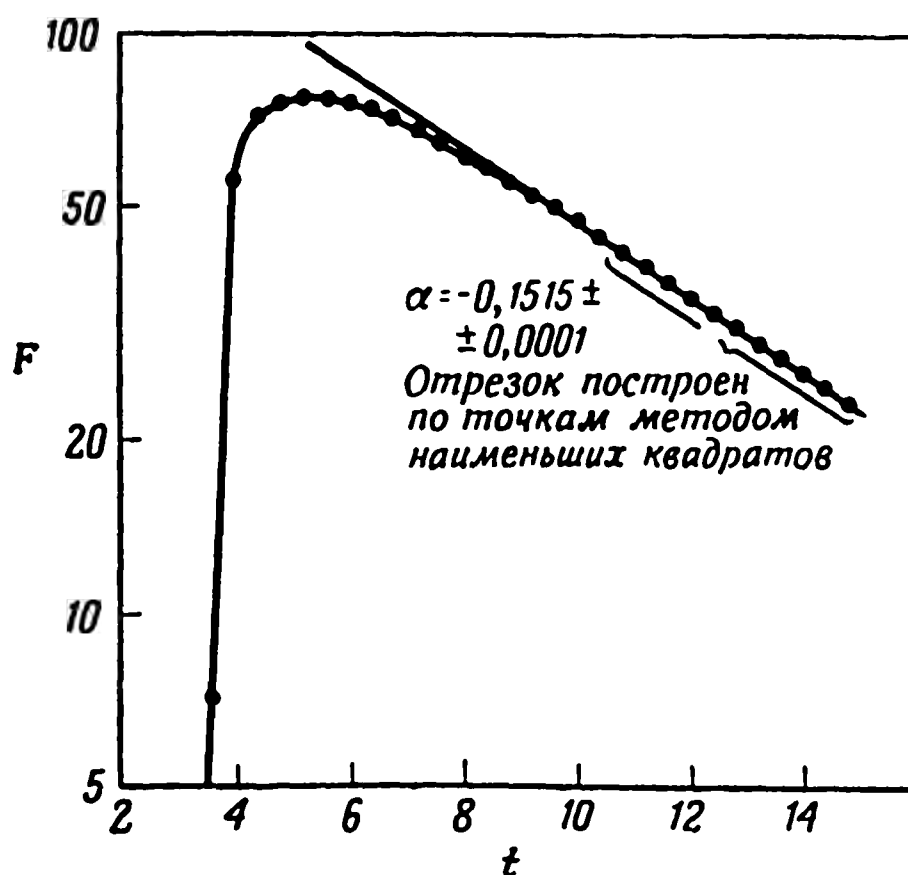
$$F = 2\pi \int_0^{\pi/2} \Psi(a \cos \theta, a \sin \theta, t) \sin \theta d\theta.$$

Поток F изображен в логарифмическом масштабе как функция от t для случая I на рис. 9.4а и для случая II на рис. 9.4б. В обоих случаях поток F практически равен нулю (как и должно быть) до момента времени $t = a/c$, когда приходит начальный импульс. Далее облако светится с интенсивностью,

которая с течением времени меняется, как это показано на рисунке. Сначала кривая F от t носит переходный характер, а за-



Р и с. 9.4а. Зависимость от времени светового потока, излучаемого облаком, радиус которого равен удвоенной средней длине свободного пробега; $I = 18$, $a = 2$.

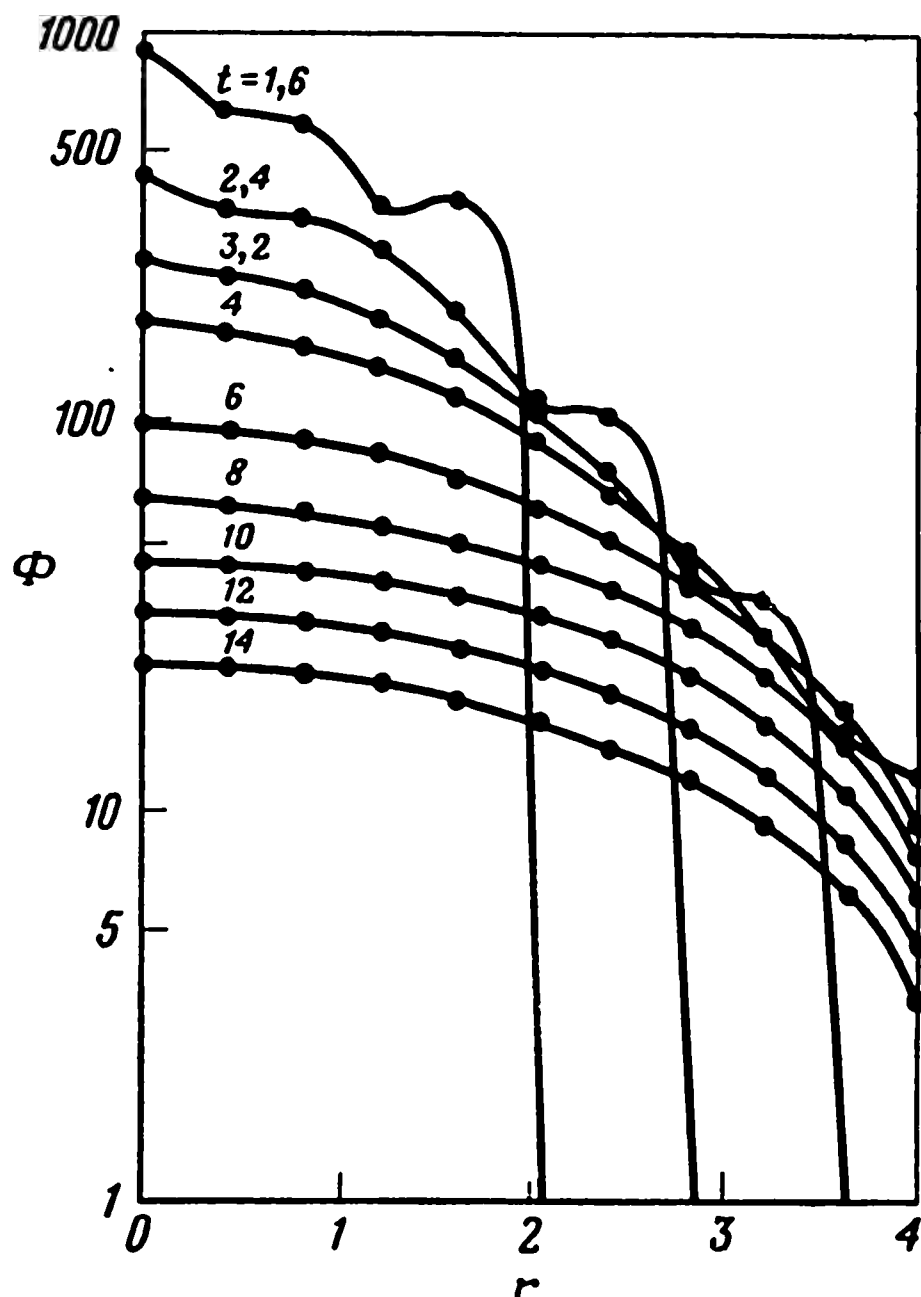


Р и с. 9.4б. То же, что на рис. 9.4а, но для облака, радиус которого равен учетверенной средней длине свободного пробега; $I = 10$, $a = 4$.

тем становится очень близкой к экспоненте с отрицательным показателем.

На рис. 9.5 показана для случая II объемная плотность $\Phi = \Phi(r, t)$ как функция от r в различные моменты времени.

На трех первых кривых ясно виден разрывный фронт излучения при $r = ct$. Колебания позади фронта, по-видимому, не соответствуют действительности, а появляются в результате погрешности аппроксимации конечноразностных уравнений. После $t = a/c$ распределение становится гладким и приближается к стационарному.



Мы не имеем точных данных для сравнения с полученными результатами, однако их точность можно оценить при помощи различных косвенных методов. Во-первых, это различные проверки, связанные с самим расчетом. Например, общее число фотонов N в обла-

Рис. 9.5. Плотность фотонов в облаке как функция радиуса в различные моменты времени.

ке в момент времени t можно получить либо как интеграл по пространству от объемной плотности

$$N = N_1 \sim \int_0^a r^2 \Phi(r) dr,$$

либо как интеграл в фазовом пространстве от плотности в нем

$$N = N_2 \sim \frac{1}{2} \int \int y \Psi(x, y) dx dy;$$

последний интеграл берется по площади всего полукруга $x^2 + y^2 \leq a^2$, $y \geq 0$. Значения N_1 и N_2 , конечно, должны быть равны. После каждого цикла машина вычисляла N_1 и N_2 (по формуле трапеций) и печатала их. В промежутке времени $0 \leq t \leq a/c$, когда фронт разрыва двигался наружу, имелось очень большое различие между N_1 и N_2 , но далее согласование было лучше. Для расчета, показанного на рис. 9.4а, разница

между N_1 и N_2 составляла приблизительно

15% для $t = a/(2c)$,

3% для $t = a/c$,

1,6% (в среднем) для $a/c < t < 2a/c$.

Если учесть, что количество точек сетки было невелико (18 на радиус), то такое расхождение не явится неожиданностью. Кроме того, есть некоторые основания считать, что даже это расхождение большей частью обусловлено неудачным масштабом (потерей значащих цифр), а не погрешностью аппроксимации, так как когда для участка $4 \leq t \leq 6$ вычисления были проведены после смены масштабов, то среднее расхождение между N_1 и N_2 на этом участке составило только 0,4%.

Другая проверка, связанная с самим расчетом: величина

$$T = N + \int F dt$$

должна быть постоянной из-за сохранения числа фотонов при $t \geq a/c$, т. е. после того, как член с источником исключается из уравнения. Если для вычисления N в последнем равенстве использовать N_2 , то величина T окажется почти постоянной; для расчета, показанного на рис. 9.4а, среднее отклонение T от ее среднего значения для $a/c < t < 2a/c$ составило 0,11%. Это, между прочим, свидетельствует о том, что N_2 (интеграл по фазовому объему) обеспечивает более точную оценку N , чем N_1 (интеграл по обычному объему).

Вторым способом проверки точности вычислений является изменение пространственной сетки. Был сделан расчет с более грубой сеткой, чем сетка, использованная при получении результатов, показанных на рис. 9.4а: $I = 12$ вместо 18. Результаты везде хорошо согласуются с полученными ранее, за исключением непосредственной окрестности момента $t = a/c$, где внезапное начало свечения было несколько менее резким, чем при $I = 18$. Для $t > a/c$ обе кривые с точностью около 0,5% имеют одинаковую форму, однако кривая для $I = 12$ всегда лежит ниже кривой для $I = 18$ примерно на 2,5%.

Еще одной оценкой точности метода является исследование асимптотического поведения решения при больших значениях t . Если мы возьмем в качестве решения уравнения (9.30) $\Psi(x, y, t) = \Psi_0(x, y)e^{\alpha t}$ (такое решение называется *собственным колебанием*), то можно показать, что соответствующая функция плотности

$$\Phi_0(r) = \alpha \frac{1+f}{2r} \int_{-r}^r \Psi_0(x, \sqrt{r^2 - x^2}) dx$$

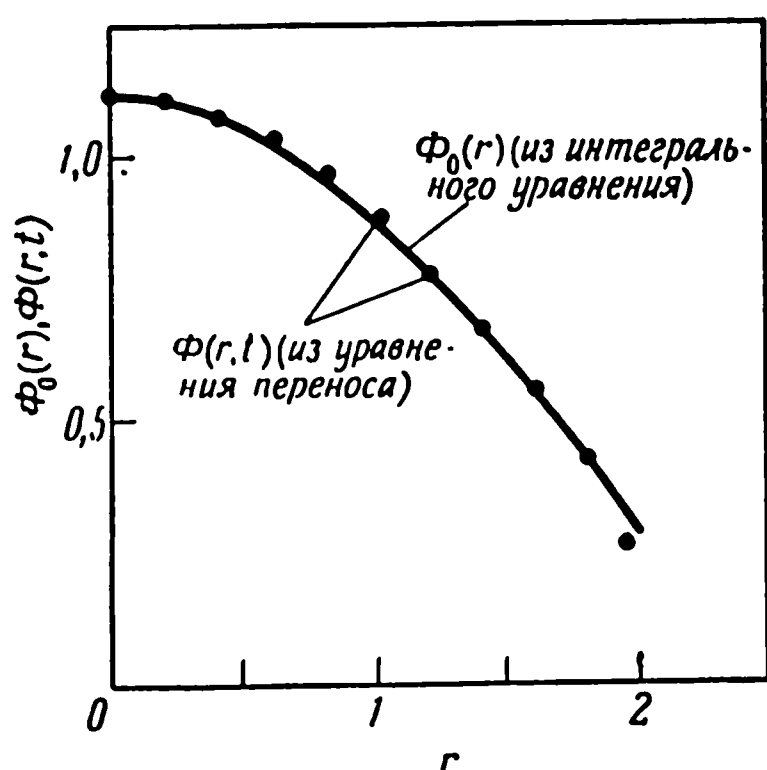
(ср. с уравнением (9.32)) является решением известного в теории переноса интегрального уравнения

$$\Phi_0(s) = \lambda \int_{-a}^a E((\sigma + \alpha/v)|s - r|) \Phi_0(r) dr, \quad (9.34)$$

где

$$\lambda = \sigma(1 + f)/2, \quad E(x) = \int_x^\infty (e^{-y} dy/y).$$

Фундаментальное решение уравнения (9.34) можно получить при помощи обычных численных методов, которые дают решение со столь малой погрешностью, что в настоящих рассуждениях оно может считаться точным.



Р и с. 9.6. Асимптотическое распределение плотности в фазовом пространстве: зависимость Φ от r при большом фиксированном значении t . Кривая соответствует точному решению интегрального уравнения (9.34).

Обычно предполагается, что общее решение уравнения переноса (9.30) представляется в виде линейной суперпозиции отдельных собственных колебаний. Для больших значений t преобладает основное собственное колебание, т. е. то, для которого значение α является наибольшим, и асимптотическая форма решения имеет вид ¹⁾

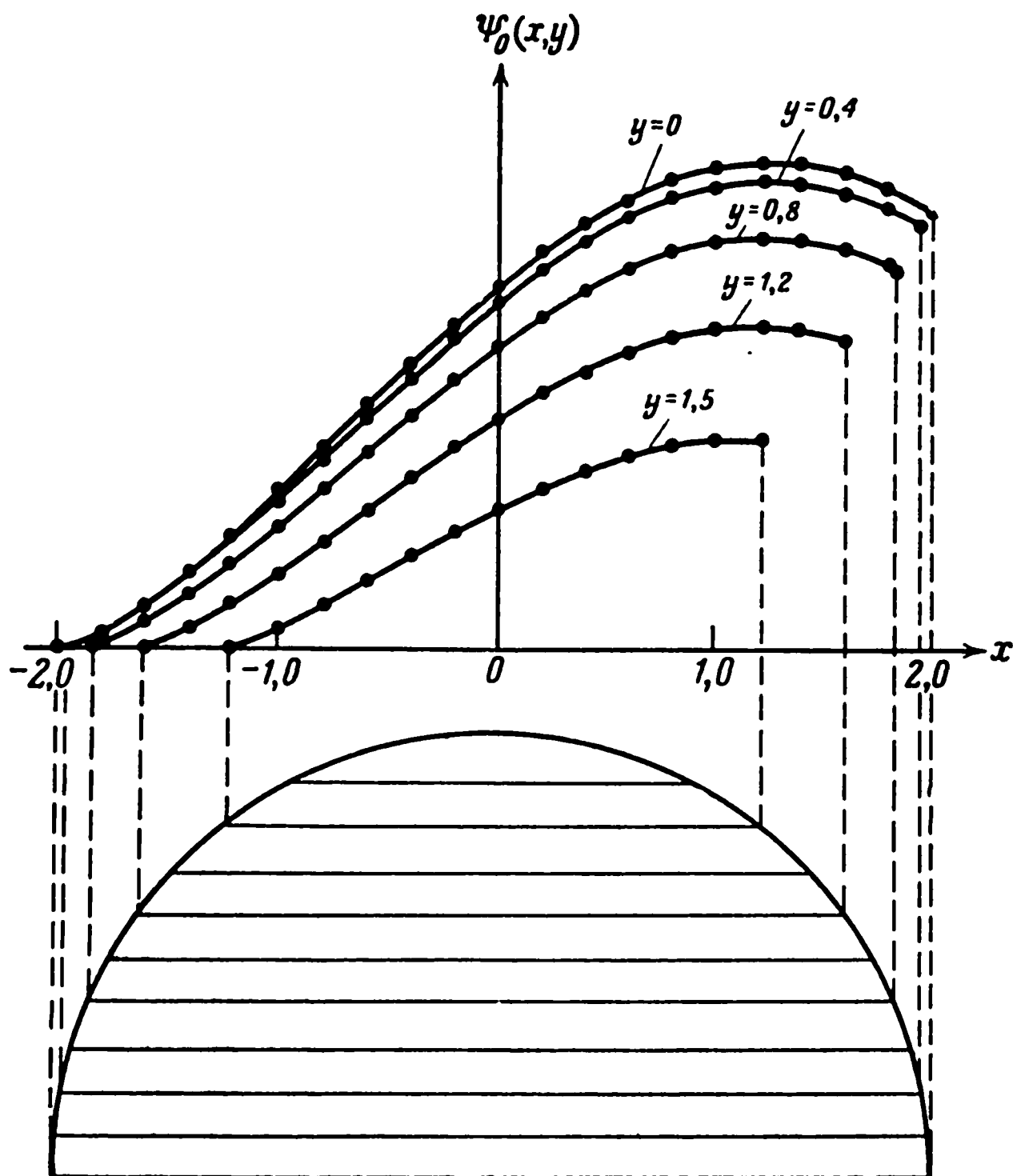
$$\Psi(x, y, t) \sim \Psi_0(x, y) e^{\alpha t}$$

и соответственно

$$\Phi(r, t) \sim \Phi_0(r) e^{\alpha t}.$$

¹⁾ В работе Ленера и Уинга [1955] показано, что в аналогичной задаче для слоя сделанное выше предположение будет ошибочным, однако выводы, основанные на нем, остаются в силе. Ленер и Уинг получили общее решение для слоя в виде линейной суперпозиции конечного числа собственных колебаний и добавочного члена, который с ростом t убывает быстрее, чем любое из собственных колебаний. Нортон [1962] показал, что для сферы существует бесконечное множество собственных колебаний.

Исследование численного решения, изображенного на рис. 9.4а, показывает, что при $4 < t < 6$ убывание является точно экспоненциальным с константой убывания $\alpha = -0,4904$. На основании этого предположим, что асимптотическая форма решения достаточно хорошо устанавливается при $t=4$, и сравним



Р и с. 9.7. Асимптотическое распределение плотности в фазовом пространстве: зависимость Ψ_0 от x при различных значениях y и при большом фиксированном значении t .

его с точным решением уравнения (9.34). Точное значение α , полученное из уравнения (9.34), равно $-0,4925$ и совпадает с приведенным выше значением с точностью до долей процента. На рис. 9.6 для одного из расчетов сравнивается зависимость $\Phi(r, t)$ от r при $t = 5$ с фундаментальным решением $\Phi_0(r)$ уравнения (9.34), нормированным надлежащим образом. Расхождение не превышает процента, за исключением точки на поверхности $r = 2$, где оно составляет 5%.

Учитывая грубость применяемой сетки, можно считать, что асимптотическое поведение нашего решения в точности таково, как требуется.

Асимптотическое распределение в фазовом пространстве, полученное численным способом, представлено на рис. 9.7, где вычерчены графики зависимости $\Psi_0(x, y)$ от x при различных значениях y .

В заключение отметим основные моменты, характерные для рассмотренного метода.

1. В координатах x, y свободный пробег нейтрона определяется условием $y = \text{const}$.

2. В пространстве переменных x, y, t выбирается прямоугольная сетка.

3. Шаг Δt выберется равным $\Delta x/v$, и уравнение интегрируется вдоль характеристик (траекторий в пространстве координат-времени).

4. Дробные интервалы сетки по x вблизи поверхности сферы специально выбираются, чтобы обеспечить точность второго порядка.

5. Член Φ , описывающий поток, центрируется по времени с помощью итерации.

6. Если источник имеет особенности, то следует по крайней мере первое рассеяние рассчитать аналитически, чтобы получить для источника более гладкое распределение.

Глава 10

ЗВУКОВЫЕ ВОЛНЫ

§ 10.1. Физические основы

В этой главе мы остановимся на задачах о малых колебаниях жидкости (газа). Вязкость предполагается пренебрежимо малой, такой же предполагается теплопроводность всюду, кроме последнего параграфа. Объемные силы, если таковые имеются, предполагаются уравновешенными градиентом статического давления.

Пусть p_0 , ρ_0 — статическое давление и плотность, и пусть $p = f(\rho)$ — уравнение состояния при адиабатических сжатии и расширении. Предположим, что $f'(\rho) > 0$, и положим $f'(\rho_0) = c^2$; величины p_0 , ρ_0 , c^2 могут зависеть от x , y , z , но не от t . Если $\mathbf{g} = \mathbf{g}(x, y, z)$ — объемная сила, отнесенная к единице объема, то мы предположим, что $\mathbf{g} - \nabla p_0 = 0$ (статическое равновесие). Если давление и плотность равны $p_0 + p$ и $\rho_0 + \rho$, где $p \ll p_0$ и $\rho \ll \rho_0$, то с точностью до величин первого порядка

$$p = c^2 \rho.$$

Если $\mathbf{v} = \mathbf{v}(x, y, z, t)$ — скорость жидкости, то уравнение движения имеет вид

$$\rho_0 \partial \mathbf{v} / \partial t = - \nabla p,$$

а уравнение неразрывности с точностью до величин первого порядка малости записывается как

$$\partial \rho / \partial t = - \rho_0 \nabla \cdot \mathbf{v}.$$

Далее,

$$\partial p / \partial t = c^2 (\partial \rho / \partial t) = - \rho_0 c^2 \nabla \cdot \mathbf{v}$$

и поэтому

$$\partial^2 p / \partial t^2 = \rho_0 c^2 \nabla \cdot (1/\rho_0) \nabla p. \quad (10.1)$$

Уравнение (10.1) — это общее линейное волновое уравнение, к которому и применяются конечноразностные методы этой главы. Но в примерах и исследовании устойчивости мы ограничимся случаем одномерного движения с постоянными ρ_0 и c^2 . В этом приближении смещение, скорость, ускорение, плотность и т. д. удовлетворяют уравнениям одного и того же вида.

Например, если $y = y(x, t)$ означает смещение (это обычный выбор зависимого переменного), то мы имеем

$$\partial^2 y / \partial t^2 = c^2 (\partial^2 y / \partial x^2), \quad c^2 = \text{const} > 0. \quad (10.2)$$

Уравнение этого вида называется *волновым уравнением*. Предположим, что заданы начальные условия

$$\begin{aligned} y(x, 0) &= f_1(x), \\ \partial y(x, 0) / \partial t &= f_2(x) \end{aligned}$$

и условие периодичности

$$y(x, t) = y(x + L, t)$$

вместо граничных условий. Для читателя будет очевидным, как распространить исследование устойчивости на многомерные случаи, если ρ_0 и c^2 остаются постоянными и периодичность сохраняется, и как обобщить конечноразностные уравнения на общий случай переменных ρ_0 и c^2 .

§ 10.2. Обычное конечноразностное уравнение

Это уравнение имеет вид

$$y_j^{n+1} - 2y_j^n + y_j^{n-1} = (c \Delta t / \Delta x)^2 (y_{j+1}^n - 2y_j^n + y_{j-1}^n), \quad (10.3)$$

и чтобы включить его в общую схему гл. 4, может показаться заманчивым положить $v = \partial y / \partial t$ и заменить уравнение (10.2) системой

$$\partial v / \partial t = c^2 (\partial^2 y / \partial x^2), \quad \partial y / \partial t = v.$$

Однако это нецелесообразно, потому что если y и v рассматривать как компоненты вектора $u(x, t)$ и взять (как это делалось в гл. 4) в качестве нормы соответствующего элемента $u \in \mathcal{R}$ величину

$$\|u\| = \left[\int_0^L (|y|^2 + |v|^2) dx \right]^{1/2},$$

то оператор $E_0(t)$ будет неограниченным и краевая задача будет поставлена некорректно. Это можно увидеть, рассматривая простые стоячие волны

$$f_1 = \sin(2\pi n x / L), \quad f_2 = 0,$$

для которых решение $u(t)$ получается легко; мы находим, что

$$\frac{\|u(t)\|}{\|u(0)\|} = \sqrt{\cos^2(2\pi n c t / L) + (2\pi n c / L)^2 \sin^2(2\pi n c t / L)},$$

но эта величина может быть сделана сколь угодно большой надлежащим выбором n и t . Этот пример показывает просто, что выбор нормы входит в понятие корректной постановки краевой задачи и что мы выбрали норму неподходящим образом.

Лучшим способом является введение вспомогательной функции $w = c \partial u / \partial x$ и замена уравнения (10.2) системой

$$\partial v / \partial t = c (\partial w / \partial x), \quad \partial w / \partial t = c (\partial v / \partial x), \quad (10.4)$$

когда в качестве нормы берется

$$\|u\| = \left[\int_0^L (|w|^2 + |v|^2) dx \right]^{1/2}.$$

Теперь оператор $E_0(t)$ ограничен: действительно, его верхняя грань равна 1, потому что норма $\|u(t)\|$ постоянна для каждого решения в силу закона сохранения энергии (квадрат этой нормы пропорционален энергии волнового движения в области $0 \leq x \leq L$).

Заметим также, что довольно очевидная разностная схема

$$\begin{aligned} (v_j^{n+1} - v_j^n) / \Delta t &= c (w_{j+1}^n - w_{j-1}^n) / (2 \Delta x), \\ (w_j^{n+1} - w_j^n) / \Delta t &= c (v_{j+1}^n - v_{j-1}^n) / (2 \Delta x) \end{aligned}$$

является неустойчивой, если отношение $\Delta t / (\Delta x)^2$ не остается ограниченным при $\Delta t, \Delta x \rightarrow 0$. Такое ограничение является практически неудобным, и мы отказываемся от этой схемы.

Чтобы избежать разностных отношений с двойным пространственным интервалом $2\Delta x$, мы будем впредь связывать значения w со средними точками интервалов и писать $w_{j+1/2}^n$ и т. д. Это только вопрос обозначений. Иногда так же трактуют и разностные отношения по времени, вводя дробные верхние индексы, в результате чего уравнения становятся более симметричными и улучшается центрирование. Однако мы будем этого избегать, чтобы как можно меньше отклоняться от обозначений, введенных в гл. 4.

Схема

$$\begin{aligned} (v_j^{n+1} - v_j^n) / \Delta t &= c (w_{j+1/2}^n - w_{j-1/2}^n) / \Delta x, \\ (w_{j-1/2}^{n+1} - w_{j-1/2}^n) / \Delta t &= c (v_j^{n+1} - v_{j-1}^{n+1}) / \Delta x \end{aligned} \quad (10.5)$$

эквивалентна обычной схеме (10.3), если отождествить

$$v_j^n \quad \text{с} \quad (y_j^n - y_j^{n-1}) / \Delta t,$$

и

$$w_{j-1/2}^n \quad \text{с} \quad c (y_j^n - y_{j-1}^n) / \Delta x,$$

Это одна из схем, исследованных Курантом, Фридрихсом и Леви [1928] в статье, в которой впервые было введено понятие устойчивости.

Матрица перехода для системы (10.5) имеет вид

$$G = G(\Delta t, k) = \begin{pmatrix} 1 & ja \\ ia & 1 - a^2 \end{pmatrix},$$

где $a = (2c \Delta t / \Delta x) \sin(k \Delta x / 2)$. Характеристическим уравнением для G является

$$\lambda^2 - \lambda(2 - a^2) + 1 = 0.$$

Если $a^2 \leq 4$, то оба корня этого уравнения по модулю равны 1. Очевидно, мы можем предположить, что

$$c \Delta t / \Delta x = \text{const} \quad \text{при} \quad \Delta t, \Delta x \rightarrow 0,$$

и тогда в силу определения a мы видим, что условием фон Неймана будет неравенство

$$\frac{c \Delta t}{\Delta x} \leq 1.$$

Достаточное условие устойчивости было получено в § 4.11. Если v_1 и v_2 — нормированные собственные векторы G , то мы находим, что $|\det(v_1, v_2)|$ ограничен снизу положительным числом для всех действительных k , и поэтому уравнения устойчивы в силу приведенной там теоремы, если $c \Delta t / \Delta x < 1$. При $c \Delta t / \Delta x = 1$ система (10.5) неустойчива; чтобы показать это, достаточно взять неограниченное решение

$$v_j^n = (-1)^{n+j} (1 - 2n), \quad w_{j+1/2}^n = (-1)^{n+j} 2n.$$

Столь медленный рост ошибок, по-видимому, не принес бы вреда на практике, но малые возмущения, как было отмечено в § 5.2, могут сделать этот рост более значительным. Однако точное равенство $c \Delta t = \Delta x$ не имеет смысла для реальных вычислений, когда коэффициент c может быть переменным и в любом случае известен только с конечным числом десятичных знаков.

§ 10.3. Неявная система

Система

$$\begin{aligned} \frac{v_j^{n+1} - v_j^n}{\Delta t} &= c \frac{w_{j+1/2}^n - w_{j-1/2}^n + w_{j+1/2}^{n+1} - w_{j-1/2}^{n+1}}{2 \Delta x}, \\ \frac{w_{j-1/2}^{n+1} - w_{j-1/2}^n}{\Delta t} &= c \frac{v_j^{n+1} - v_{j-1}^{n+1} + v_j^n - v_{j-1}^n}{2 \Delta x}, \end{aligned} \quad (10.6)$$

эквивалентная система второго порядка

$$\frac{y_j^{n+1} - 2y_j^n + y_j^{n-1}}{(\Delta t)^2} = c^2 \frac{(\delta^2 y)_j^{n+1} + 2(\delta^2 y)_j^n + (\delta^2 y)_j^{n-1}}{4(\Delta x)^2}, \quad (10.7)$$

имеет матрицу перехода

$$G = \begin{pmatrix} \frac{1 - a^2/4}{1 + a^2/4} & \frac{ia}{1 + a^2/4} \\ \frac{ia}{1 + a^2/4} & \frac{1 - a^2/4}{1 + a^2/4} \end{pmatrix}.$$

Оба собственных значения G по модулю равны 1, а G^*G — единичная матрица, так что система (10.6) безусловно устойчива.

Из уравнения (10.7), очевидно, следует, что решение этой неявной системы может быть получено с помощью алгоритма, данного в гл. 8 для решения неявных систем, аппроксимирующих уравнение диффузии. Поэтому уравнения (10.6) образуют практически удовлетворительную систему для волнового уравнения.

§ 10.4. Одновременное распространение звука и тепла

В потоке сжимаемой жидкости часто наблюдается заметное изменение температуры от точки к точке, и перенос энергии посредством теплопроводности может оказать значительное влияние на движение. Тогда параболическое уравнение теплопроводности присоединяется к гиперболическим уравнениям движения жидкости и оба процесса должны рассчитываться одновременно. Это же имеет место и при бесконечно малых или акустических колебаниях и обуславливает поглощение ультразвуковых волн.

Пусть давление, удельный объем и удельная внутренняя энергия суть $p_0 + p$, $V_0 + V$ и $\mathcal{E}_0 + \mathcal{E}$, где p_0 , V_0 и $\mathcal{E}_0$ — невозмущенные величины и $p \ll p_0$, $V \ll V_0$, $\mathcal{E} \ll \mathcal{E}_0$, и пусть u — скорость частиц. Величины p , V , $\mathcal{E}$ и u являются функциями от x и t , а $c = \sqrt{p_0 V_0}$ представляет собой изотермическую скорость звука. Будем считать, что уравнение состояния имеет вид

$$\mathcal{E}_0 + \mathcal{E} = (p_0 + p)(V_0 + V)/(\gamma - 1),$$

и обозначим через σ отношение коэффициента теплопроводности к удельной теплоемкости при постоянном объеме.

Если ввести вспомогательные зависимые переменные $w = cV/V_0$ и $e = \mathcal{E}/c$, то дифференциальные уравнения с точностью

до малых величин первого порядка примут вид

$$\begin{aligned}\frac{\partial u}{\partial t} &= c \frac{\partial}{\partial x} [\omega - (\gamma - 1) e], \\ \frac{\partial \omega}{\partial t} &= c \frac{\partial u}{\partial x}, \\ \frac{\partial e}{\partial t} &= \sigma \frac{\partial^2 e}{\partial x^2} - c \frac{\partial u}{\partial x}.\end{aligned}\tag{10.8}$$

Конечноразностные уравнения для этой системы можно строить различными способами. Простой явной системой является

$$\begin{aligned}\frac{u_j^{n+1} - u_j^n}{\Delta t} &= c \frac{\omega_{j+1/2}^n - \omega_{j-1/2}^n - (\gamma - 1)(e_{j+1/2}^n - e_{j-1/2}^n)}{\Delta x}, \\ \frac{\omega_{j+1/2}^{n+1} - \omega_{j+1/2}^n}{\Delta t} &= c \frac{u_{j+1}^{n+1} - u_j^{n+1}}{\Delta x}, \\ \frac{e_{j+1/2}^{n+1} - e_{j+1/2}^n}{\Delta t} &= \sigma \frac{e_{j+3/2}^n - 2e_{j+1/2}^n + e_{j-1/2}^n}{(\Delta x)^2} - c \frac{u_{j+1}^{n+1} - u_j^{n+1}}{\Delta x}.\end{aligned}\tag{10.9}$$

Значения u из верхнего слоя использованы во втором и в третьем уравнениях с той же целью, что и во втором уравнении системы (10.5) при отсутствии теплопроводности (напомним, что это было необходимо для получения приемлемого условия устойчивости в задаче о звуковых волнах). Несмотря на это система (10.9) будет в действительности явной, потому что сначала можно решить первое уравнение и найти значения u_j^{n+1} и u_{j+1}^{n+1} , необходимые для двух других уравнений.

Эта система оказалась удовлетворительной. Если звуковые волны и распространение тепла разъединены, то соответствующими условиями устойчивости являются неравенства

$$\sqrt{\gamma} c \Delta t / \Delta x < 1, \quad \sigma \Delta t / (\Delta x)^2 < 1/2.\tag{10.10}$$

Конечно, эти условия являются необходимыми. В пределе, когда $\Delta t \rightarrow 0$ и $\Delta x \rightarrow 0$, выполнение второго из этих условий влечет за собой выполнение первого, и можно вообще предположить, что именно оно и является условием устойчивости. В реальных вычислениях нужно, конечно, выбирать Δx и Δt так, чтобы они удовлетворяли как первому, так и второму условию.

Чтобы избежать малости шага по времени, требуемой условиями (10.10), часто применяют неявную схему третьего дифференциального уравнения; например,

$$\frac{e_{j+1/2}^{n+1} - e_{j+1/2}^n}{\Delta t} = \sigma \frac{e_{j+3/2}^{n+1} - 2e_{j+1/2}^{n+1} + e_{j-1/2}^{n+1}}{(\Delta x)^2} - c \frac{u_{j+1}^{n+1} - u_j^{n+1}}{\Delta x},\tag{10.11}$$

первые же два уравнения системы (10.9) остаются без изменения.

В первом издании книги предполагалось, что для этой системы условием устойчивости должно быть неравенство $c\Delta t/\Delta x < 1$. При этом использовалась изотермическая скорость звука c , потому что при неустойчивости обычно возрастают возмущения с очень малой длиной волны, а для них время установления теплового равновесия от вершины до впадины меньше, чем период волны. Это время пропорционально квадрату длины волны λ ; сжатия и разрежения в волновом движении при $\lambda \ll \sigma/c$ в гораздо большей степени являются изотермическими, чем адиабатическими.

Правильность этого предположения подтвердил Моримото [1962]. Однако в пределе при $\sigma \rightarrow 0$ условием устойчивости будет более сильное неравенство $\sqrt{\gamma} c \Delta t/\Delta x < 1$, и это наводит на мысль о том, что в случае конечной сетки, когда нельзя считать, что $\Delta x \ll \sigma/c$, существует практическое условие устойчивости, в которое входят две безразмерные постоянные

$$\nu = c \Delta t/\Delta x, \quad \mu = \sigma \Delta t/(\Delta x)^2 \quad (10.12)$$

и γ ; вероятно оно должно быть промежуточным между приведенными выше условиями, а именно должно иметь вид $\nu < \nu_0$, где ν_0 есть некоторая функция от γ и μ , значения которой лежат между $1/\sqrt{\gamma}$ и 1.

Чтобы отыскать это практическое условие устойчивости, воспользуемся сформулированным в § 9.7 положением о том, что степень роста компонент Фурье приближенного решения не должна превосходить максимально возможной степени роста точного решения. Так как амплитуды звуковых волн в точном решении не возрастают (в действительности для рассматриваемой задачи благодаря действию теплопроводности они убывают, но для длинных волн это убывание незначительно), для множителя перехода g вместо условия $|g| \leq 1 + O(\Delta t)$ должно быть $|g| \leq 1$. Уравнение для g имеет вид

$$\begin{vmatrix} g - 1 & -2i\nu \sin \alpha & 2i\nu (\gamma - 1) \sin \alpha \\ -2ig\nu \sin \alpha & g - 1 & 0 \\ 2ig\nu \sin \alpha & 0 & g - 1 + 2\mu (1 - \cos 2\alpha) g \end{vmatrix} = 0,$$

где $\alpha = (1/2)k\Delta x$. Требование, что $|g| \leq 1$ для всех α , приводит к неравенству

$$\nu < \sqrt{(1 + 2\mu)/(\gamma + 2\mu)}, \quad (10.13)$$

которое является практическим условием устойчивости; в § 6.4 было показано, что это условие является также достаточным

для устойчивости в том смысле, что если оно выполняется, то некоторая положительно определенная квадратичная форма от u, w, e и $\Delta_+ e_j = e_{j+1} - e_j$ остается ограниченной на всем временном интервале.

Как и предполагалось, неравенство (10.13) показывает, что устойчивость определяется адиабатической скоростью звука при $\mu \ll 1$, изотермической скоростью звука при $\mu \gg 1$ и промежуточной скоростью при $\mu \sim 1$.

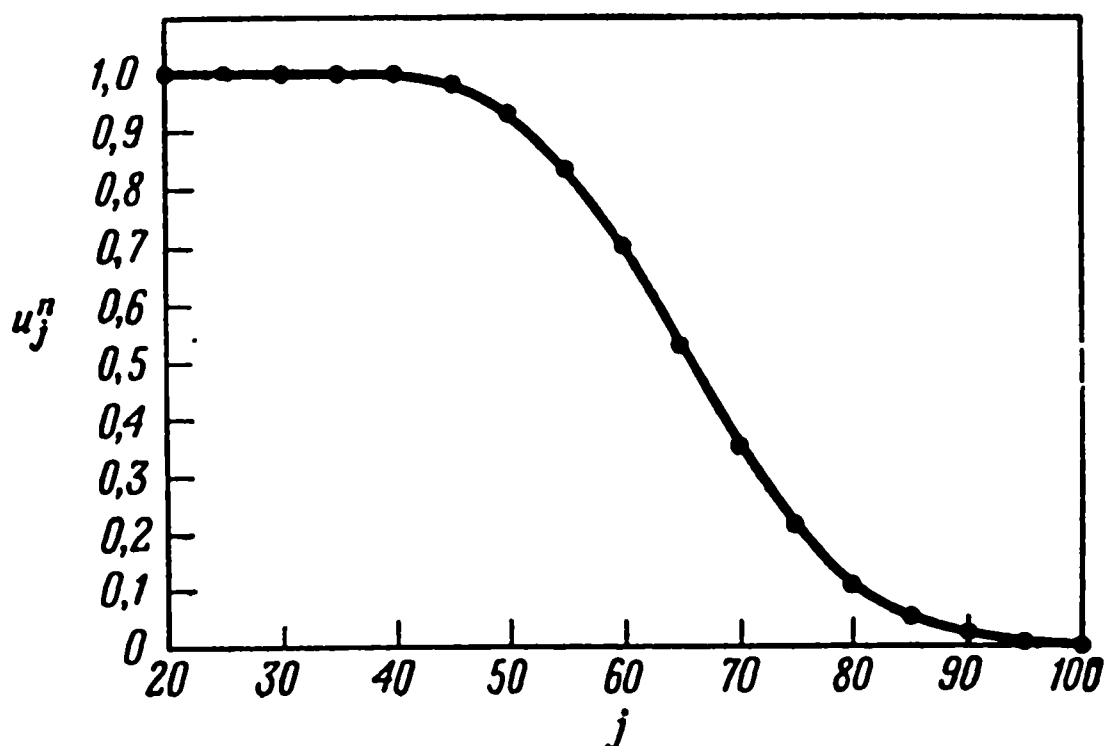


Рис. 10.1. Профиль скорости, рассчитанный для разрывной в начале звуковой волны при наличии теплопроводности; 165 шагов по времени; $\gamma = 3,0$; $\mu = 1,1547$; $\nu = 0,57735 = c \Delta t / \Delta x$.

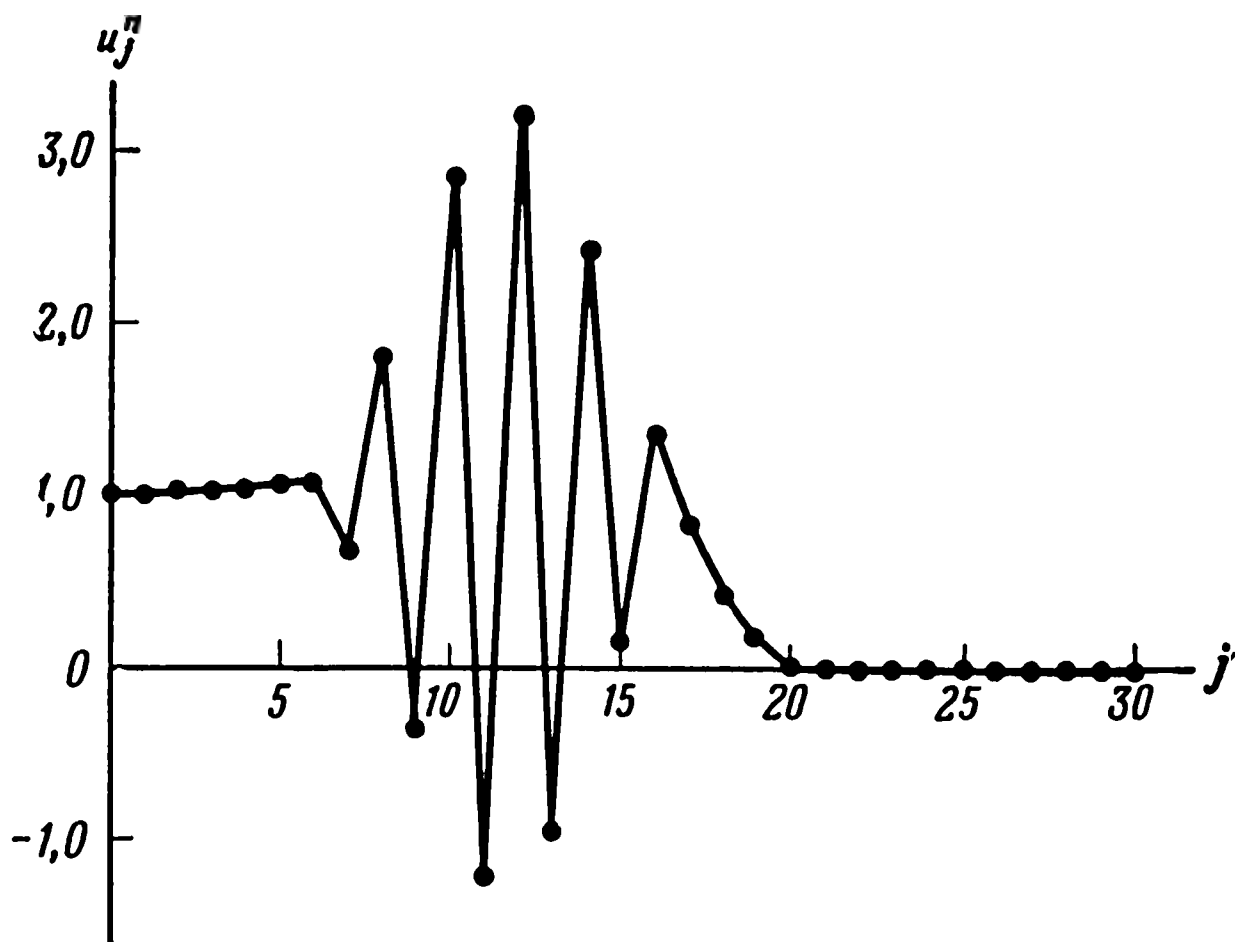
В одной серии вычислений разностные уравнения (первые два из системы (10.9) и уравнение (10.11)) были применены для расчета распространения волны, которая задавалась как простой скачок или бесконечно малый удар (бесконечно малый в том смысле, что применимы линеаризованные уравнения. Такая волна распространяется со скоростью $\sqrt{\gamma}c$, а профиль ее постепенно размазывается таким образом, что напоминает интегральную кривую нормального распределения

$$\int_{-\infty}^x \exp[-y^2/2] dy$$

и зона размазывания увеличивается как $\sqrt{t}$. Для этого случая известно также приближенное (но достаточно хорошее) аналитическое решение.

На рис. 10.1 показан профиль, получившийся в устойчивом расчете после 165 шагов по времени, а на рис. 10.2 приведены значения, полученные в неустойчивом расчете всего лишь после пяти шагов по времени; в обоих расчетах было выполнено обычное условие устойчивости $c\Delta t / \Delta x < 1$. Рис. 10.3 обобщает ре-

зультаты всей серии расчетов: каждая точка указывает значения μ и ν , для которых расчет оказался устойчивым, а каждый

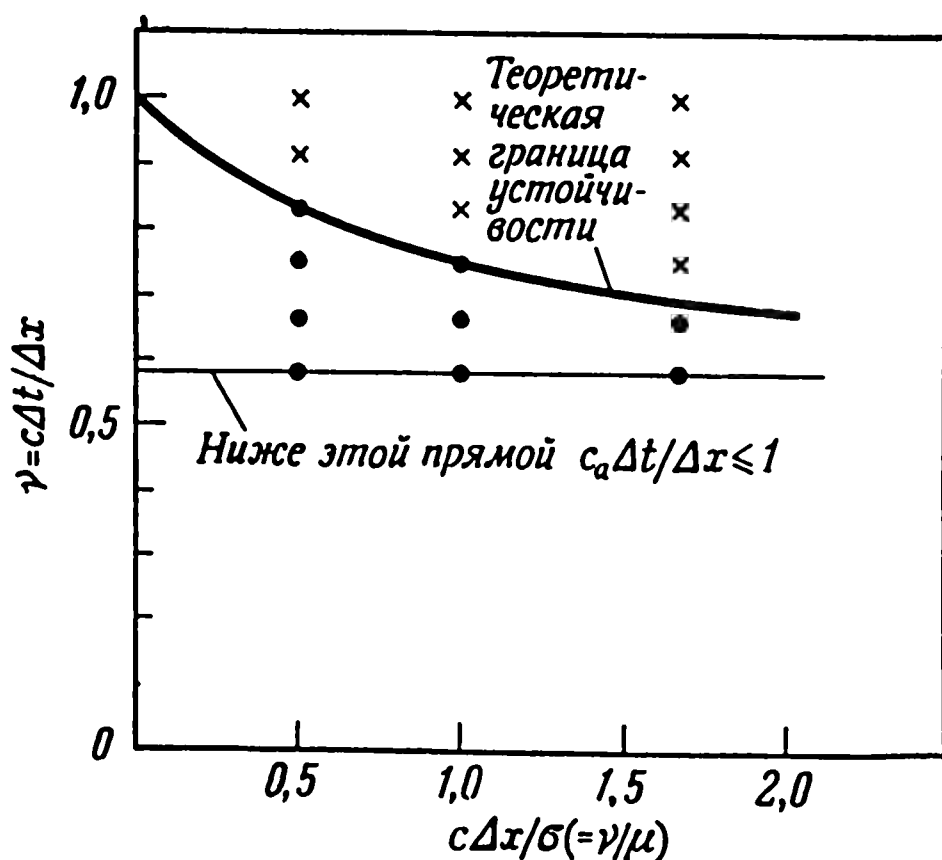


Р и с. 10.2. Тот же профиль; 5 шагов по времени; практический критерий устойчивости (10.13) не выполняется, хотя величина $c \Delta t / \Delta x$ и не превосходит единицы; $\gamma = 3,0$; $\mu = 0,447846$; $\nu = 0,746410 = c \Delta t / \Delta x$.

крестик указывает значения, для которых расчет оказался неустойчивым; кривая описывает максимально допустимые значения ν как функцию от ν/μ , согласно условию (10.13). Видно, что

Р и с. 10.3. Диаграмма устойчивости для задачи об одновременном распространении звука и тепла при $\gamma = 3$.

При расчетах вариант считался устойчивым, если после примерно 100 шагов по времени профиль скорости еще оставался таким же гладким, как на рис. 10.1, и неустойчивым, если значение $|u_j^n|$ в какой-либо точке превосходило 3,0 (если такое превышение имело место, то оно всегда проявлялось на первых пяти шагах по времени); c — изометрическая скорость звука, c_a — адиабатическая скорость звука, $\bullet$ — устойчивый вариант, $\times$ — неустойчивый вариант.



наблюдаемая в расчетах устойчивость согласуется с полученной теоретически. Расчеты, отвечающие двум точкам на кривой, для которых неравенство (10.13) заменяется точным равенством, также оказались устойчивыми.

Чтобы понять сущность такого типа неустойчивости, рассмотрим расчет с $\nu = c\Delta t/\Delta x = 0,9$ и $\gamma = 3$. Согласно обычному критерию, это — устойчивая схема, и при $\Delta t \rightarrow 0$ имеет место сходимость. Однако если Δt таково, что μ также равно 0,9 (именно этот случай был выбран для иллюстрации), то, согласно приведенному выше характеристическому уравнению, $g \approx -2,6$ для компоненты Фурье с $k\Delta x = \pi$; следовательно, эта компонента возрастает в тысячу раз менее чем за 10 шагов по времени. Чтобы добиться устойчивости при условии $c\Delta t/\Delta x = 0,9$, Δx и Δt необходимо умножить примерно на 0,25, так что останется в силе неравенство (10.13).

Другой подход состоит в том, что сохраняется явный характер разностных уравнений, но третье уравнение решается с более мелким шагом по времени, чем первые два, например с шагом $\Delta t/K$, где K — натуральное число. Тогда каждый шаг решения гидродинамических уравнений сопровождается K шагами решения уравнения теплопроводности. Этот алгоритм можно записать с помощью дробных верхних индексов:

$$\frac{e_{j+1/2}^{n+(m+1)/K} - e_{j+1/2}^{n+m/K}}{\Delta t/K} = \sigma \frac{e_{j+1/2}^{n+m/K} - 2e_{j+1/2}^{n+m/K} + e_{j-1/2}^{n+m/K}}{(\Delta x)^2} - c \frac{u_{j+1}^{n+1} - u_j^{n+1}}{\Delta x},$$

$$m = 0, 1, \dots, K - 1.$$

Предполагается, что в этом случае условие устойчивости определяется неравенствами

$$\frac{c \Delta t}{\Delta x} < 1 \quad \text{и} \quad \frac{\sigma \Delta t/K}{(\Delta x)^2} < 1/2.$$

§ 10.5. Практический критерий устойчивости

Приведенный выше пример и подобные ему в § 9.7 и 11.6 показывают, что устойчивость, как она обычно определяется, не всегда равнозначна практической ценности метода. Из устойчивости следует сходимость в пределе при $\Delta t \rightarrow 0$, если выполнены предположения теоремы Лакса об эквивалентности, но может случиться, что несмотря на это для конечных значений Δt , используемых на практике, ошибки неприемлемо возрастают. Чтобы ни одна из компонент Фурье приближенного решения не могла расти быстрее, чем максимально растущее точное решение, необходимо наложить некоторое ограничение на константу, характеризующую член $O(\Delta t)$ в условии фон Неймана $|g| \leq 1 + O(\Delta t)$.

Согласно § 4.3 компонента Фурье точного решения имеет вид

$$\hat{u}(k, t) = e^{tP(ik)} \hat{u}(k, 0), \quad (10.14)$$

где $P = P(q) = P(q_1, \dots, q_d)$ — введенная там же матрица, элементы которой являются полиномами от $q_i = \partial/\partial x_i$ и которая определяет дифференциальный оператор A системы уравнений $du/dt = Au$. Пусть $m = \text{Max}[\text{Re } \lambda]$ (λ — собственное значение матрицы $P(ik)$), где максимум берется по всем действительным векторам k и всем собственным значениям для каждого k ; тогда решение не может расти быстрее, чем функция e^{mt} , возможно, умноженная на полином от t . Поэтому для собственных значений матрицы перехода G мы возьмем в качестве *модифицированного условия фон Неймана* неравенство

$$|g| = |g(\Delta t, k)| \leq e^{m \Delta t}.$$

Можно спросить, не является ли условие (10.15) слишком строгим и не исключает ли оно для некоторых задач все возможные схемы, но это представляется маловероятным в силу теоремы Крайса о диссипативных схемах для симметричных гиперболических систем (§ 5.4), где $m = 0$ и где в действительности требовалось, чтобы $|g| < 1$ для всех k , кроме $k = 0$. В самом деле, вполне может быть, что критерием приемлемости схемы порядка точности $2r - 1$ является условие

$$|g| \leq e^{m \Delta t} (1 - \delta |\xi|^{2r})$$

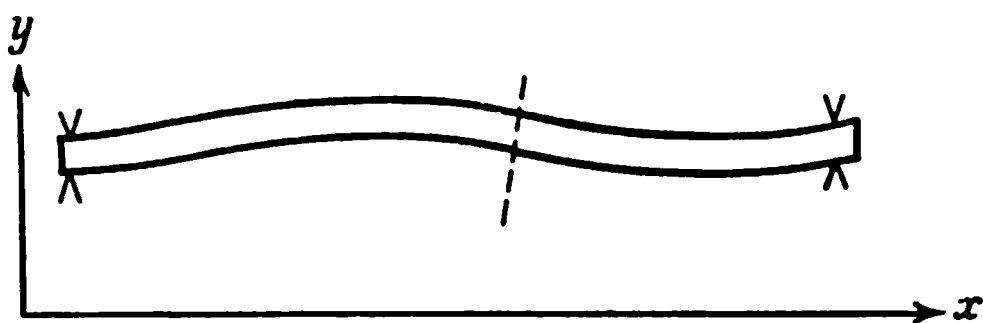
для некоторого $\delta > 0$ и для всех векторов $\xi = (k_1 \Delta x_1, \dots, k_d \Delta x_d)$, компоненты которых заключены между $-\pi$ и π . Недостаточность обычной теории в некотором смысле состоит в том, что она не определяет значений ни m , ни δ . Теперь мы определили m , а δ , вероятно, всегда должно быть величиной порядка $1/2 \sqrt{d} \pi^{-2r}$.

Глава 11

УПРУГИЕ КОЛЕБАНИЯ

§ 11.1. Поперечные колебания тонкого стержня

Задачи динамической теории упругости ставят перед численным анализом некоторые интересные вопросы, которые мы продемонстрируем на простых одномерных задачах. Предположим, что прямой тонкий стержень (или балка или струна) длины L с опертymi или заделанными концами приведен в колебание. Допустим, что деформации остаются все время малыми, так что



Р и с. 11.1. Деформированный упругий стержень с опертymi концами и воображаемое сечение, рассматриваемое в тексте.

применим закон Гука. Для этого требуется, чтобы радиус кривизны кривой, по которой изогнут стержень, был значительно больше, чем радиус поперечного сечения стержня. Для струны это не означает ее близости к прямой: струну рояля можно завязать узлом, не переходя предел упругости. Чтобы линеаризовать задачу, мы, однако, будем в дальнейшем предполагать, что стержень остается близким к прямой и что амплитуда колебаний мала. Мы будем также предполагать, что колебания являются поперечными, и обозначим через $y = y(x, t)$ смещение от положения равновесия на расстоянии x от одного из концов стержня. В этих обозначениях предположения, ведущие к линейности, записываются так: $|y(x, t)| \ll L$, $|\partial y(x, t)/\partial x| \ll 1$.

Если стержень разбить на две части воображаемым сечением, показанным на рис. 11.1 штриховой линией, то левая часть будет действовать на правую с некоторой силой F и крутящим моментом M . Мы предполагаем пока, что в стержне отсутствуют продольные усилия; тогда сила F является поперечной. Если выбрать положительное направление F вдоль оси y , а положительное направление M против часовой стрелки, то упругие свойства стержня можно выразить уравнениями

$$M = -\kappa \frac{\partial^2 y}{\partial x^2}, \quad F = \frac{\partial}{\partial x} \kappa \frac{\partial^2 y}{\partial x^2},$$

где κ — положительная величина, зависящая от материала и поперечного сечения стержня. Момент и сила, с которыми правая часть действует на левую, очевидно, равны $-M$ и $-F$. Поэтому сила, действующая на отрезок длины δl , равна величине $-\delta l \partial F / \partial x$; следовательно, уравнение движения имеет вид

$$\mu \frac{\partial^2 y}{\partial t^2} = - \frac{\partial^2}{\partial x^2} \kappa \frac{\partial^2 y}{\partial x^2},$$

где $\mu = \mu(x)$ — линейная плотность стержня.

Мы дополнительно упростим задачу, предполагая, что κ и μ постоянны. Тогда уравнение принимает вид

$$\frac{\partial^2 y}{\partial t^2} = - a^2 \frac{\partial^4 y}{\partial x^4}, \quad a = \sqrt{\kappa/\mu} = \text{const.} \quad (11.1)$$

Типичным граничным условием является условие «свободной опоры», согласно которому опоры препятствуют поперечному смещению, но не влияют на вращающие моменты, т. е.

$$y = 0, \quad \frac{\partial^2 y}{\partial x^2} = 0 \quad \text{при} \quad x = 0 \quad \text{и} \quad x = L. \quad (11.2)$$

Чтобы переписать уравнение (11.1) как два уравнения первого порядка по времени, положим $v = \partial y / \partial t$ и $w = a \partial^2 y / \partial x^2$; тогда

$$\frac{\partial v}{\partial t} = - a \frac{\partial^2 w}{\partial x^2}, \quad \frac{\partial w}{\partial t} = a \frac{\partial^2 v}{\partial x^2}. \quad (11.3)$$

Разностные уравнения, которые мы будем рассматривать, могут быть применены либо к уравнению (11.1), либо к системе (11.3).

Заметим попутно, что вследствие характера граничных условий (11.2) они могут быть заменены условиями периодичности и нечетности, а именно

$$y(x, t) = y(x + 2L, t) = -y(-x, t), \quad (11.2a)$$

для функций, которые теперь определены для всех действительных x . Это доказывается следующим образом. Дифференциальные уравнения (11.3) инвариантны относительно преобразований $x + 2L \rightarrow x$ и $-x \rightarrow x$. Поэтому за начальные значения $v(x, 0)$ и $w(x, 0)$ можно принять два любых нечетных тригонометрических полинома с периодом $2L$, и полученное при этом точное решение будет также инвариантно относительно этих преобразований. Таким образом, единственное решение $y(x, t)$ уравнения (11.1), удовлетворяющее условиям (11.2a), определяется через $a \partial^2 y / \partial x^2 = w$. Эта функция y в силу симметрии удовлетворяет также условиям (11.2). Иначе говоря, две крайние задачи (с граничными условиями (11.2) и с граничными

условиями (11.2a)) обладают разрешающими операторами $E_0(t)$, которые совпадают на нечетных тригонометрических полиномах. Теперь обе краевые задачи поставлены корректно (мы опускаем доказательство) в банаховом пространстве $\mathcal{B}$, состоящем из пар функций v, w на отрезке $0 \leq x \leq L$ с нормой $\left[\int_0^L (|v|^2 + |w|^2) dx \right]^{1/2}$,

так что $E_0(t)$ ограничены для обеих задач. Далее, нечетные тригонометрические полиномы образуют плотное в $\mathcal{B}$ множество, и поэтому обобщенные разрешающие операторы $E(t)$ для обеих задач идентичны, что и требовалось доказать.

Свойства простых конечноразностных систем для этой задачи столь сходны со свойствами систем для волнового уравнения, приведенных в предыдущей главе, что мы остановимся на них весьма кратко.

§ 11.2. Явные разностные уравнения

Если применять разности по времени вперед в обоих уравнениях (11.3), то получающаяся разностная система будет неустойчивой, если Δt не стремится к нулю как $(\Delta x)^3$, и поэтому является неудовлетворительной.

Если брать одну разность по времени вперед, а другую назад, то мы получим систему

$$\begin{aligned} \frac{v_j^{n+1} - v_j^n}{\Delta t} &= -a \frac{(\delta^2 w)_j^n}{(\Delta x)^2}, \\ \frac{w_j^{n+1} - w_j^n}{\Delta t} &= a \frac{(\delta^2 v)_j^{n+1}}{(\Delta x)^2}, \end{aligned} \quad (11.4)$$

которая эквивалентна разностному уравнению

$$\frac{y_j^{n+1} - 2y_j^n + y_j^{n-1}}{(\Delta t)^2} = -a^2 \frac{(\delta^4 y)_j^n}{(\Delta x)^4}, \quad (11.5)$$

составленному для уравнения (11.1), и является явной. Погрешность аппроксимации для уравнения (11.5) равна $O((\Delta t)^2) + O((\Delta x)^2)$.

Если для краткости ввести обозначение

$$[4a \Delta t / (\Delta x)^2] \sin^2(k \Delta x / 2) = \omega,$$

то матрица перехода для системы (11.4) будет иметь вид

$$G = G(\Delta t, k) = \begin{pmatrix} 1 & \omega \\ -\omega & 1 - \omega^2 \end{pmatrix}.$$

Собственные значения матрицы G равны по модулю единице при $|\omega| < 2$, следовательно, если мы предположим, что $\Delta t \rightarrow 0$ как $(\Delta x)^2$, то необходимым для устойчивости условием фон Неймана будет неравенство

$$a \Delta t / (\Delta x)^2 \leq 1/2.$$

Достаточное для устойчивости условие может быть получено из доказанной в § 4.11 теоремы точно так же, как и для соответствующей системы, аппроксимирующей волновое уравнение. Действительно, с точностью до обозначений указанная выше матрица G с помощью унитарного преобразования приводится к виду (4.20). Если v_1 и v_2 — нормированные собственные векторы матрицы G , то нетрудно показать, что $|\det(v_1, v_2)|^2 = 1 - (\omega^2/4)$; эта величина ограничена снизу положительной постоянной для всех действительных k , если

$$a \Delta t / (\Delta x)^2 < 1/2;$$

следовательно, это неравенство является достаточным условием устойчивости для системы (11.4)

§ 11.3. Неявная система

Из многих мыслимых неявных систем мы рассмотрим одну, аналогичную системе (10.6) для волнового уравнения, а именно систему

$$\begin{aligned} \frac{v_j^{n+1} - v_j^n}{\Delta t} &= -a \frac{(\delta^2 w)_j^{n+1} + (\delta^2 w)_j^n}{2(\Delta x)^2}, \\ \frac{w_j^{n+1} - w_j^n}{\Delta t} &= a \frac{(\delta^2 v)_j^{n+1} + (\delta^2 v)_j^n}{2(\Delta x)^2}. \end{aligned} \quad (11.6)$$

Матрица перехода для этой системы имеет вид

$$G = \begin{bmatrix} \frac{1 - \omega^2/4}{1 + \omega^2/4} & \frac{\omega}{1 + \omega^2/4} \\ \frac{-\omega}{1 + \omega^2/4} & \frac{1 - \omega^2/4}{1 + \omega^2/4} \end{bmatrix}$$

и, как нетрудно видеть, является унитарной. Поэтому 1) собственные значения G лежат на единичной окружности, и условие фон Неймана всегда выполнено, 2) матрица G нормальна, и условие фон Неймана достаточно для устойчивости. Следовательно, система (11.6) всегда устойчива. Погрешность аппроксимации эквивалентного уравнения второго порядка (по

времени) для функции y равна $O((\Delta t)^2) + O((\Delta x)^2)$, так же как для явной системы, исследованной в предыдущем параграфе.

§ 11.4. Достоинства неявной системы

Задача о поперечных колебаниях стержня является примером задачи, в которой применение неявной системы действительно оправдано. Погрешность аппроксимации имеет одинаковый порядок относительно Δt и Δx как для системы (11.6), так и для системы (11.4). Поэтому, если требуется повышенная точность, то с точки зрения точности нужно уменьшать Δt в той же мере, что и Δx ; условие же устойчивости для явной системы требует, чтобы Δt убывало как $(\Delta x)^2$. Это, очевидно, приводит к применению значительно большего числа шагов по времени, чем это диктуется соображениями одной лишь точности. Далее, может возникнуть опасность, что разности, которые нужно прибавлять в v^n и w^n , чтобы получить v^{n+1} и w^{n+1} , могут получиться столь же малыми, как и возмущения, возникающие за счет ошибок округления. Неявная система свободна от этого недостатка.

§ 11.5. Решение неявных уравнений произвольного порядка

Алгоритм для решения уравнений неявной системы (11.6) аналогичен таковому для неявных систем, встречающихся в задачах теплопроводности и распространения звуковых волн, но в то же время система (11.6) обладает достаточным количеством новых свойств, что оправдывает ее рассмотрение. Если составить эквивалентную систему для y_j^{n+1} , то теперь в матрице коэффициентов системы не обращающиеся в нуль элементы будут расположены не на трех, а на пяти диагоналях.

Типичное уравнение совместной системы может быть записано как пятичленное рекуррентное соотношение для y_j^{n+1} или два трехчленных рекуррентных соотношения для v_j^{n+1} и w_j^{n+1} . Мы возьмем последнюю интерпретацию и введем двумерный вектор

$$u_j = \begin{pmatrix} v_j^{n+1} \\ w_j^{n+1} \end{pmatrix}.$$

Тогда из системы (11.6) легко усмотреть, что рекуррентное соотношение (после умножения на Δt) имеет вид

$$-A_j u_{j+1} + B_j u_j - C_j u_{j-1} = d_j, \quad j = 1, 2, \dots, J-1 \quad (11.7)$$

(отметим сходство с уравнением (8.15) в исследовании уравнения диффузии), где A_j , B_j и C_j — матрицы второго порядка

$$B_j = \begin{bmatrix} 1 & -\frac{a \Delta t}{(\Delta x)^2} \\ \frac{a \Delta t}{(\Delta x)^2} & 1 \end{bmatrix},$$

$$A_j = C_j = \begin{bmatrix} 0 & -\frac{1}{2} \frac{a \Delta t}{(\Delta x)^2} \\ \frac{1}{2} \frac{a \Delta t}{(\Delta x)^2} & 0 \end{bmatrix},$$

а d_j — заданный вектор. Заметим, что в данной задаче A_j , B_j и C_j не зависят от j и что $A = C$, $B = I + A + C$.

Чтобы получить алгоритм, применимый к более общим задачам, мы не будем использовать этих специальных свойств данных матриц. Этот алгоритм можно будет применять также к задачам, в которых вектор u_j имеет любое число компонент. Предположим что матрицы A_j , B_j и C_j невырождены.

Граничные условия на свободно опертых концах стержня дают равенства $u_0 = u_J = 0$, но мы допустим, что заданы более общие граничные условия, связывающие u и $\partial u / \partial x$ на каждом конце, а именно

$$u_0 = H u_1 + l, \quad (11.8)$$

$$u_{J-1} = M u_J + n, \quad (11.9)$$

где H и M — некоторые заданные матрицы, а l и n — заданные векторы. Граничные условия еще более общего вида будут рассмотрены ниже.

Теперь система (11.7) и условие (11.8) определяют двухпараметрическое семейство¹⁾ решений системы (11.7). Если дано некоторое множество невырожденных матриц E_j и векторов f_j , то уравнения

$$u_j = E_j u_{j+1} + f_j, \quad j = 0, 1, \dots, \quad (11.10)$$

также имеют двухпараметрическое семейство решений; эти два семейства будут идентичны, если

$$\begin{aligned} E_0 &= H, \quad f_0 = l, \\ E_j &= (B_j - C_j E_{j-1})^{-1} A_j, \quad j \geq 1, \\ f_j &= (B_j - C_j E_{j-1})^{-1} (d_j - C_j f_{j-1}). \end{aligned} \quad (11.11)$$

¹⁾ В более общем случае k -параметрическое семейство, если мы имеем дело с k -мерными векторами и матрицами порядка k .

Последние два соотношения получены подстановкой $E_{j-1}u_j + f_{j-1}$ вместо u_{j-1} в систему (11.7) и решением полученных уравнений относительно u_j , а затем отождествлением коэффициентов в найденном выражении для u_j с аналогичными им в уравнении (11.10).

Как и в задаче для уравнения диффузии, уравнения (11.11) служат для определения E_j и f_j по индукции ($j = 0, 1, 2, \dots$). Индукция может продолжаться (столь долго, сколь долго матрица $(B_{j+1} - C_{j+1}E_j)$ остается невырожденной, а мы считаем, что это условие все время выполняется) до тех пор, пока не будут вычислены E_j и f_j для $j = 1, 2, \dots, J-1$. Заметим, что E_j , получаемые таким образом, невырождены. При вычислении E_j нужно обращать матрицу второго порядка. Это можно сделать либо с помощью заранее заготовленных формул, либо путем обращения к имеющейся в машине специальной подпрограмме; если число компонент больше двух, то последнее, вероятно, предпочтительнее.

Векторы u_j получаются теперь по индукции для убывающего j . Сначала исключением u_{J-1} из уравнений

$$u_{J-1} = E_{J-1}u_J + f_{J-1},$$

$$u_{J-1} = Mu_J + n$$

определяется u_J . Мы должны предположить, что матрица $E_{J-1} - M$ невырожденная. Мы делаем это предположение потому, что в противном случае решение либо не существует, либо не единственно, а это значит, что граничные условия в конечноразностной форме (11.8), (11.9) были сформулированы некорректно. Затем из уравнений (11.10) решение определяется по индукции в направлении убывания j ($j = J-1, J-2, \dots, 0$).

Чтобы эта схема была приемлемой, элементы матриц E_j и компоненты векторов f_j должны оставаться в пределах числовой шкалы машины. Мы покажем сначала, что при некоторых предположениях матрицы E_j будут равномерно ограничены. Замечая, что для матриц обращение произведения равно произведению обращенных сомножителей в обратном порядке, мы можем записать второе из соотношений (11.11) в виде

$$\begin{aligned} E_j &= [B_j B_j^{-1} (B_j - C_j E_{j-1})]^{-1} A_j = [B_j^{-1} (B_j - C_j E_{j-1})]^{-1} B_j^{-1} A_j = \\ &= [I - B_j^{-1} C_j E_{j-1}]^{-1} B_j^{-1} A_j. \end{aligned}$$

Обозначив через $\|M\|$ норму матрицы M , используем то обстоя-

тельство¹⁾, что если $\|M\| < 1$, то $\|(I - M)^{-1}\| \leq (1 - \|M\|)^{-1}$, и что $\|MN\| \leq \|M\| \cdot \|N\|$. Применяя эти соотношения к приведенной выше рекуррентной формуле для E_j и полагая $\|B_j^{-1}A_j\| = \gamma_j$, $\|B_j^{-1}C_j\| = \delta_j$, находим, что если $\delta_j\|E_{j-1}\| < 1$, то

$$\|E_j\| \leq \frac{\gamma_j}{1 - \delta_j\|E_{j-1}\|}. \quad (11.12)$$

Это наводит на мысль рассмотреть последовательность $\xi_0, \xi_1, \dots$, заданную с помощью соотношений

$$\xi_0 = \|E_0\| = \|H\|, \quad \xi_j = \gamma_j / (1 - \delta_j \xi_{j-1}), \quad (11.13)$$

так как по индукции тогда получается, что если

$$\delta_j \xi_{j-1} < 1 \quad \text{при } j = 1, 2, \dots, \quad (11.14)$$

то

$$\|E_j\| \leq \xi_j \quad \text{при } j = 0, 1, \dots \quad (11.15)$$

Поэтому мы заинтересованы в том, чтобы найти такие условия, наложенные на ξ_0 и на последовательности (γ_j) , (δ_j) , чтобы 1) последовательность (ξ_j) была ограниченной, 2) неравенство $\delta_j \xi_{j-1} < 1$ выполнялось бы для всех j . Рассмотрим простой случай, когда γ_j и δ_j не зависят от j , как это имеет место для задач с постоянными коэффициентами. Положим $\gamma_j = \gamma$, $\delta_j = \delta$ и заметим, что $\gamma > 0$, $\delta > 0$.

На рис. 11.2 изображен график величины

$$\xi' = \gamma / (1 - \delta \xi)$$

как функции от ξ . Преобразование

$$\gamma / (1 - \delta \xi) \rightarrow \xi$$

получается перемещением в направлении стрелок на чертеже. Если кривая $\xi' = \xi'(\xi)$ пересекает биссектрису координатного угла (как это показано на чертеже), то каждая точка ξ , лежащая левее точки A , перемещается при этом преобразовании ближе к точке B . Условием для такого пересечения является

¹⁾ Чтобы это доказать, заметим, во-первых, что $I - M$ имеет обратную матрицу, потому что если бы матрица $I - M$ была вырожденной, то мы имели бы $\det(I - M) = 0$, и одно из собственных значений M было бы равно 1, что противоречит условию $\|M\| < 1$. Во-вторых, если $v = (I - M)w$, то

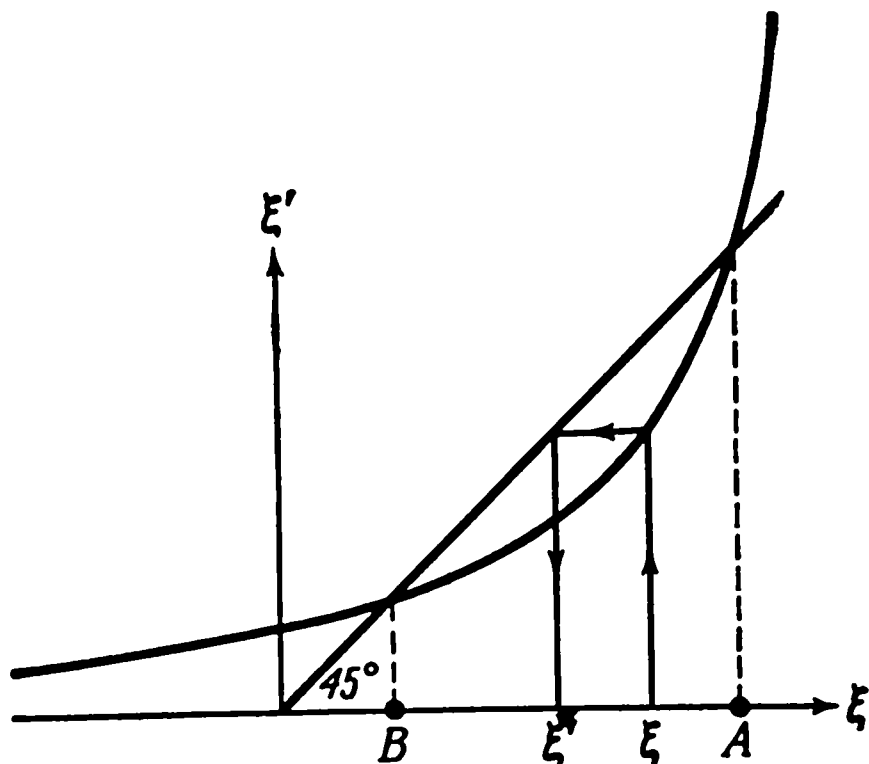
$$\begin{aligned} \text{Max} \frac{\|(I - M)^{-1} v\|}{\|v\|} &= \text{Max} \frac{\|w\|}{\|(I - M)w\|} \leq \frac{\|w\|}{\|w\| - \|Mw\|} \leq \\ &\leq \frac{\|w\|}{\|w\| - \|M\| \|w\|} = \frac{1}{1 - \|M\|}, \end{aligned}$$

потому что $\|Mw\| \leq \|M\| \cdot \|w\| \leq \|w\|$.

неравенство $\gamma\delta < 1/4$, и поэтому $A = (1 + \sqrt{1 - 4\gamma\delta})/2\delta$. Очевидно, что если $\xi < A$, то $\delta\xi < 1$. Поэтому все условия будут выполнены, если

$$\gamma\delta < 1/4 \quad \text{и} \quad \xi_0 < (1 + \sqrt{1 - 4\gamma\delta})/2\delta. \quad (11.16)$$

Если взять матрицы A , B и C такими, как в задаче об однородном свободно опертом стержне, то условия (11.16) будут выполнены, потому что $A = C$ и, следовательно, $\gamma = \delta$; в частности $\gamma = \delta = \|B^{-1}A\|$. Нормой $B^{-1}A$ является квадратный ко-



Р и с. 11.2. Построение для исследования ограниченности матриц E_j , определяемых вторым равенством (11.11) (см. также равенства (11.13)).

рень из модуля наибольшего собственного значения матрицы $(B^{-1}A)^*(B^{-1}A)$, который, как легко показать, будет равен

$$\gamma = \delta = \frac{1}{2} \frac{a \Delta t / (\Delta x)^2}{\sqrt{1 + [a \Delta t / (\Delta x)^2]^2}} < 1/2; \quad (11.17)$$

кроме того, граничным условием в случае свободной опоры будет $0 = H = E_0$, и поэтому $\xi_0 = 0$. Отсюда следует, что матрицы E_j остаются ограниченными при неограниченно возрастающем j .

Можно коснуться вопроса и о граничных условиях более общего вида. Пусть p будет число компонент вектора u , так что A_j , B_j и C_j являются матрицами порядка p . В уравнениях (11.7) число неизвестных превосходит число уравнений на $2p$. Допустим, что половина необходимых дополнительных уравнений (т. е. p из них) является линейными условиями на левом конце, а вторая половина — линейными условиями на правом конце. При этом мы предполагаем, что половина этих условий содержит только компоненты векторов $u_0, u_1, \dots, u_{j_0}$, а вторая половина — только компоненты векторов $u_J, u_{J-1}, \dots, u_{J-j_0}$, где j_0 — в общем случае небольшое натуральное число, которое остается фиксированным при $\Delta x \rightarrow 0$ и, следовательно, при $J \rightarrow \infty$,

так как $\Delta x = L/J$. Тогда в пределе эти условия превращаются в граничные условия в том смысле, что они содержат только значения функций и некоторых их производных при $x = 0$ или при $x = L$. Предположим, что $J > 2j_0$.

Граничные условия будем предполагать независимыми; уравнения (11.7) с граничными условиями на левом конце стержня имеют общее решение, линейно зависящее от p параметров, которые мы обозначим через $\lambda_1, \dots, \lambda_p$. В частности, компоненты u_j для каждого j , т. е. величины u_{ji} ($i = 1, \dots, p$), являются линейными функциями этих параметров. Обозначим через U_j матрицу, заданную соотношениями

$$(U_j)_{jk} = \partial u_{ji} / \partial \lambda_k,$$

так что

$$u_j = U_j \lambda + g_j, \quad j = 0, 1, \dots, \quad (11.18)$$

где λ — вектор

$$\lambda = \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_p \end{pmatrix},$$

а g_j — вектор, компоненты которого не зависят от этих параметров.

Чтобы интерпретировать эти формулы для некоторых частных случаев, заметим, что если граничное условие на левом конце имеет вид (11.8), как это предполагалось в предыдущих рассуждениях, то мы можем принять параметры $\lambda_1, \dots, \lambda_p$ за компоненты вектора u_1 , так что U_1 будет единичной матрицей, которая является (помимо всего прочего) невырожденной. Однако граничные условия для стержня с заделанными концами, имеющие вид $y = 0$ и $dy/dx = 0$ на концах, при переводе на конечноразностный язык можно записать как $v_0 = v_1 = 0$, так что за такие параметры могут быть приняты w_0 и w_1 . В этом случае первые три матрицы U_j будут иметь вид

$$U_0 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad U_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad U_2 = \begin{pmatrix} 0 & 2(\Delta x)^2/a \Delta t \\ -1 & 2 \end{pmatrix},$$

и первой невырожденной матрицей является U_2 . (Для получения U_2 используется уравнение (11.7) с $j = 1$.) Граничное условие на свободном конце стержня, т. е. условие отсутствия поперечной силы и крутящего момента, или $\partial^2 y / \partial x^2 = 0$, $\partial^3 y / \partial x^3 = 0$, принимает вид $w_0 = w_1 = 0$. При этом получается тот же результат, что и ранее, за исключением того, что v и w меняются ролями.

Пусть теперь U_{j_1} будет первой невырожденной матрицей в последовательности $U_1, U_2, \dots$ (заметим, что U_0 в нее не включается). Тогда

$$\lambda = U_{j_1}^{-1} (\mathbf{u}_{j_1} - \mathbf{g}_{j_1}).$$

Следовательно,

$$\mathbf{u}_{j_1-1} = U_{j_1-1} U_{j_1}^{-1} (\mathbf{u}_{j_1} - \mathbf{g}_{j_1}) + \mathbf{g}_{j_1-1}.$$

Это можно записать в виде равенства

$$\mathbf{u}_{j_1-1} = H \mathbf{u}_{j_1} + \mathbf{l}, \quad (11.19)$$

которое имеет такую же форму, как и условие (11.8), за исключением того, что оно применяется в точке $j = j_1$, а не в точке $j = 1$. Вычисление E_j и $\mathbf{f}_j$ по индукции проводится так же, как и раньше, но начинается с $j = j_1$. Граничное условие на правом конце может быть аналогично взято в форме (11.9), скажем в точке $j = J - j_2$, и вычисление $\mathbf{u}_j$ по индукции для уменьшающегося j должно начинаться с $j = J - j_2$. После этого остаются лишь немногие значения $\mathbf{u}_j$, которые должны быть вычислены вблизи каждого из концов — они могут быть получены из уравнений (11.7).

На практике $H, \mathbf{l}, M$ и $\mathbf{n}$ обычно весьма просто выражаются через первоначальные граничные условия и, быть может, один или два коэффициента системы (11.7). По-видимому, обычно лучше всего сначала выписать эти выражения, а затем вставить их непосредственно в программу.

Другое замечание состоит в том, что на практике идеализированный случай, когда γ_j и δ_j не зависят от j , вообще не имеет места; вместо этого точки A и B на рис. 11.2 могут слегка перемещаться при переходе от одной итерации преобразования

$$(1 - \delta_j \xi) / \gamma_j \rightarrow \xi$$

к другой. Но точка B является точкой положительной устойчивости в том смысле, что каждая точка ξ , лежащая левее точки A , движется определенно к B с каждой итерацией, и интуитивно ясно, что если начальное значение $\xi = \xi_0$ лежит заметно левее точки A , а точки A и B при каждой итерации перемещаются достаточно мало и остаются на некотором расстоянии друг от друга, то ξ также будет оставаться слева от A . В данном случае аналитическими методами трудно определить заранее, выполняются ли эти условия, и при проведении расчета следует полагаться либо на машину с большим диапазоном допустимых числовых значений, либо на специальную подпрограмму, которая бы автоматически реагировала надлежащим образом,

когда матрица E_j выходит из допустимой числовой шкалы, т. е. когда матрица $B_j - C_j E_{j-1}$ стремится стать вырожденной.

§ 11.6. Колебания продольно напряженного стержня

Если в предыдущие рассуждения включить член, соответствующий продольному напряжению, то появятся два новых обстоятельства, заслуживающих внимания.

Если T означает продольное напряжение, то уравнение движения принимает вид

$$\mu \frac{\partial^2 y}{\partial t^2} = -\kappa \frac{\partial^4 y}{\partial x^4} + T \frac{\partial^2 y}{\partial x^2}$$

и вместо системы (11.3) мы получим систему

$$\frac{\partial v}{\partial t} = -a \frac{\partial^2 w}{\partial x^2} + b \frac{\partial w}{\partial x}, \quad \frac{\partial w}{\partial t} = +a \frac{\partial^2 v}{\partial x^2} + b \frac{\partial v}{\partial x}, \quad (11.20)$$

где $v = \partial y / \partial t$ и $w = a \partial^2 y / \partial x^2 + b \partial y / \partial x$, причем $a = \sqrt{\kappa / \mu}$ и $b = \sqrt{T / \mu}$.

Неявная разностная система

$$v_j^{n+1} - v_j^n = -a \frac{(\delta^2 w)_j^{n+1} + (\delta^2 w)_j^n}{2(\Delta x)^2} + b \frac{w_{j+1}^{n+1} - w_{j-1}^{n+1} + w_{j+1}^n - w_{j-1}^n}{4\Delta x}, \quad (11.21)$$

$$\frac{w_j^{n+1} - w_j^n}{\Delta t} = a \frac{(\delta^2 v)_j^{n+1} + (\delta^2 v)_j^n}{2(\Delta x)^2} + b \frac{v_{j+1}^{n+1} - v_{j-1}^{n+1} + v_{j+1}^n - v_{j-1}^n}{4\Delta x},$$

аналогичная системе (11.6), безусловно устойчива.

Однако в данном случае могут возникнуть трудности в применении алгоритма, данного для решения неявных уравнений в предыдущем параграфе. Матрицы A_j , B_j и C_j уравнения (11.7) за счет дополнительного члена изменяются так, что теперь вместо соотношения (11.17) мы имеем

$$\gamma = \delta = \frac{1}{2} \frac{a \Delta t / (\Delta x)^2 + b \Delta t / (2 \Delta x)}{\sqrt{1 + (a \Delta t / (\Delta x)^2)^2}}$$

и условие $\gamma \delta < 1/4$ (см. неравенства (11.16)) может выполняться или не выполняться в зависимости от значений a , b , Δt , Δx . Если $\gamma \delta > 1/4$, то на рис. 11.2 не будет пересечения, последовательность ξ_j не будет ограничена и матрицы E_j , вероятно, выйдут из шкалы. В этом случае *устойчивость разностных уравнений достигается за счет устойчивости этого частного алгоритма для их решения.*

Это затруднение существенно усугубляется, если характер колебаний обусловлен в большей мере натяжением, нежели работой на изгиб (как у струны рояля), потому что тогда $b = \sqrt{T/\mu}$ велико, а $a = \sqrt{\kappa/\mu}$ мало. Тогда эта задача качественно эквивалентна задаче о колебаниях струны, для которой уравнением движения является одномерное волновое уравнение. Это наводит на мысль рассмотреть явные разностные системы, аналогичные тем, которые оказались пригодными для волнового уравнения. Одной из таких систем является

$$\begin{aligned} \frac{v_j^{n+1} - v_j^n}{\Delta t} &= -a \frac{(\delta^2 w)_j^n}{(\Delta x)^2} + b \frac{w_{j+1}^n - w_{j-1}^n}{2 \Delta x}, \\ \frac{w_j^{n+1} - w_j^n}{\Delta t} &= a \frac{(\delta^2 v)_j^{n+1}}{(\Delta x)^2} + b \frac{v_{j+1}^{n+1} - v_{j-1}^{n+1}}{2 \Delta x}. \end{aligned} \quad (11.22)$$

Она содержит разность вперед по времени в первом из уравнений и разность назад по времени во втором и аналогична системе (10.5) для волнового уравнения и системе (11.4) для колебаний ненапряженного стержня.

Условием устойчивости для системы (11.22), как легко показать, будет неравенство

$$a \Delta t / (\Delta x)^2 < 1/2, \quad (11.23)$$

которое на практике не слишком обременительно, потому что при разумно выбранной величине Δx в силу малости a величина Δt не очень сильно ограничивается этим неравенством. В действительности здесь имеется опасность того, что Δt может оказаться недостаточно ограниченным неравенством (11.23), а это указывает на основную слабость понятия устойчивости, подобную той, о которой упоминалось в § 9.7 в связи с уравнением переноса и в § 10.4 в связи с задачей об одновременном распространении звука и тепла. Чтобы увидеть эту опасность, заметим, что если в системе (11.20) полностью опустить члены, содержащие a , то для системы (11.22) получится следующее условие устойчивости:

$$b \Delta t / (2 \Delta x) < 1. \quad (11.24)$$

В пределе при $\Delta t \rightarrow 0$ и $\Delta x \rightarrow 0$ выполнение условия (11.23) влечет за собой, конечно, выполнение условия (11.24), каковы бы ни были значения a и b , если $a \neq 0$. Но если колебания определяются в большей степени натяжением, чем работой на изгиб, так что b велико, а a мало, то действительная точечная сетка, используемая при вычислениях, может быть достаточно мелкой и удовлетворять условию (11.23), но тем не менее не будет удовлетворять условию (11.24) с большим запасом. Интуитивно ясно, что тогда в решении могут развиваться весьма

интенсивные колебания, как в случае нарушения условия устойчивости для волнового уравнения. Как в парадоксе, упомянутом в гл. 9, программист хочет не просто знать, что будет в пределе при Δt и $\Delta x \rightarrow a$, а хочет знать, что будет происходить для точечной сетки, которой он пользуется в действительности, потому что одной устойчивости в пределе ему недостаточно.

Чтобы получить практический критерий для конечноразностной схемы (11.22), мы используем (как это было сделано в гл. 9 и 10) положение о том, что скорость роста амплитуды любой компоненты Фурье не будет ощутимо большей, чем максимальная скорость роста, ожидаемая (и предполагающаяся грубо известной) для точного решения. Но для стержня со свободно опертыми концами (как напряженного, так и ненапряженного) имеет место закон сохранения энергии, согласно которому амплитуда каждой компоненты остается постоянной во времени. Попробуем поэтому потребовать, чтобы для конечноразностных уравнений амплитуды также не возрастали бы со временем. (Это не всегда возможно даже для консервативных систем, но оказывается возможным здесь.) Это означает, что в то время как условие фон Неймана, вообще говоря, допускает $\lambda \leq 1 + O(\Delta t)$, мы будем здесь требовать, чтобы выполнялось неравенство $\lambda \leq 1$.

Для системы (11.22) матрица перехода имеет вид

$$G = G(\Delta t, k) = \begin{pmatrix} 1 & \beta \\ -\bar{\beta} & 1 - |\beta|^2 \end{pmatrix},$$

где

$$\beta = \frac{4a \Delta t}{(\Delta x)^2} \sin^2 \frac{k \Delta x}{2} + i \frac{b \Delta t}{\Delta x} \sin k \Delta x.$$

Характеристическим уравнением для G является

$$\lambda^2 - \lambda(2 - |\beta|^2) + 1 = 0.$$

Мы хотим, чтобы корни этого уравнения лежали внутри единичного круга или на его границе, следовательно, должно выполняться неравенство $|\beta|^2 \leq 4$, а не только неравенство $|\beta|^2 \leq 4 + O(\Delta t)$, как раньше. Мы требуем поэтому, чтобы для всех действительных k

$$\left(\frac{4a \Delta t}{(\Delta x)^2} \right)^2 \sin^4 \frac{k \Delta x}{2} + \left(\frac{b \Delta t}{\Delta x} \right)^2 \sin^2 k \Delta x \leq 4. \quad (11.25)$$

Мажорируя левую часть этого неравенства путем замены синусов на единицы, мы получаем для системы (11.22) простое достаточное условие устойчивости

$$\left(\frac{2a \Delta t}{(\Delta x)^2} \right)^2 + \left(\frac{b \Delta t}{2 \Delta x} \right)^2 < 1, \quad (11.26)$$

которое является вполне подходящим для большинства приложений; заметим, что в нем содержится как условие (11.23), так и условие (11.24).

Интересно сравнить это с условиями устойчивости, получающимися для этой системы энергетическими методами гл. 6, так как это один из простейших примеров, в которых энергетический метод практически не позволяет получить весь интервал устойчивости, определяемый методом Фурье. Если обозначить $2a\Delta t/(\Delta x)^2$ через μ , а $b\Delta t/2\Delta x$ через ν , то уравнения (11.22) можно переписать так:

$$\begin{aligned} v^{n+1} - v^n &= -\frac{1}{2}\mu\delta^2 w^n + 2\nu\Delta_0 w^n \\ w^{n+1} - w^n &= \frac{1}{2}\mu\delta^2 v^{n+1} + 2\nu\Delta_0 v^{n+1}; \end{aligned}$$

здесь $\Delta_0 u_j = \frac{1}{2}(u_{j+1} - u_{j-1})$. Умножив первое из этих уравнений $v^{n+1} + v^n$, а второе $w^{n+1} + w^n$, сложив их вместе и просуммировав по всем точкам сетки, получим

$$\begin{aligned} \|v^{n+1}\|^2 + \|w^{n+1}\|^2 - \|v^n\|^2 - \|w^n\|^2 &= \\ &= \frac{1}{2}\mu[(w^{n+1} + w^n, \delta^2 v^{n+1}) - (v^{n+1} + v^n, \delta^2 w^n)] + \\ &\quad + 2\nu[(w^{n+1} + w^n, \Delta_0 v^{n+1}) + (v^{n+1} + v^n, \Delta_0 w^n)]. \end{aligned}$$

В силу формул суммирования по частям (6.17) — (6.19)

$$(w^{n+1} + w^n, \delta^2 v^{n+1}) = -(\Delta_+(w^{n+1} + w^n), \Delta_+ v^{n+1}),$$

и аналогично также для второго члена правой части

$$(w^{n+1} + w^n, \Delta_0 v^{n+1}) = -(\Delta_0(w^{n+1} + w^n), v^{n+1}).$$

Объединяя эти члены, мы находим, что указанное выше уравнение приводится к виду $S_{n+1} = S_n$, где

$$S_n = \|v^n\|^2 + \|w^n\|^2 + \frac{1}{2}\mu(\Delta_+ v^n, \Delta_+ w^n) + 2\nu(v^n, \Delta_0 w^n).$$

Следовательно, устойчивость имеет место, если $S_n \geq \geq \text{const}(\|v^n\|^2 + \|w^n\|^2)$. Грубая оценка скалярных произведений в S_n дает

$$|(\Delta_+ v^n, \Delta_+ w^n)| \leq 2(\|v^n\|^2 + \|w^n\|^2)$$

и

$$|(v^n, \Delta_0 w^n)| \leq \frac{1}{2}(\|v^n\|^2 + \|w^n\|^2),$$

что приводит к достаточному условию устойчивости

$$\mu + \nu < 1; \tag{11.27}$$

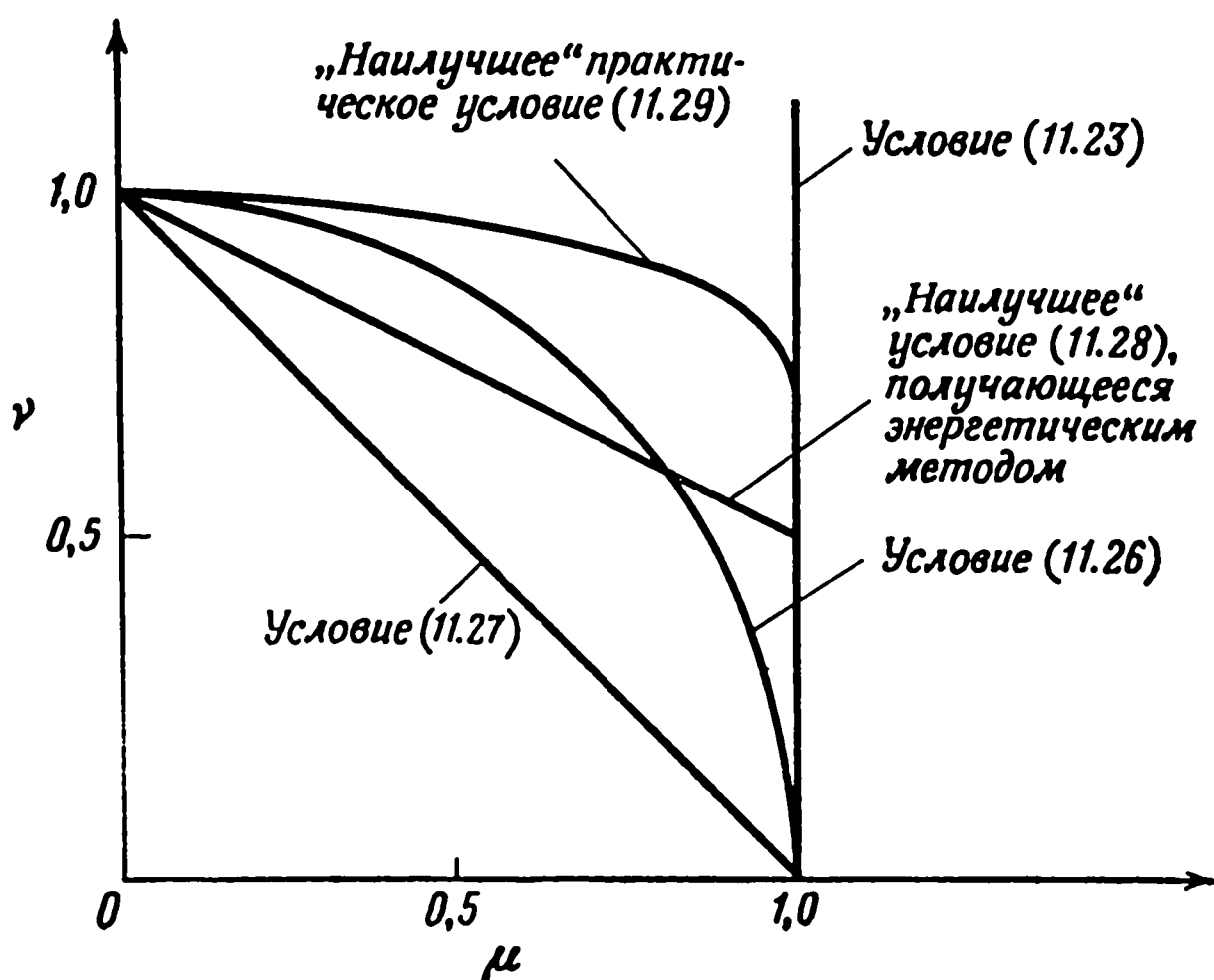
на рис. 11.3 оно сравнивается с условием (11.26). Однако при более точной оценке этих скалярных произведений, полагая $(\Delta_+ v^n, \Delta_+ w^n) = -(v^n, \delta^2 w^n)$, будем иметь

$$\begin{aligned} | -\frac{1}{2}\mu (v^n, \delta^2 w^n) + 2\nu (v^n, \Delta_0 w^n) | = \\ = | \sum_{(j)} v_j^n [(\nu - \frac{1}{2}\mu) w_{j+1}^n + \mu w_j^n - (\nu + \frac{1}{2}\mu) w_{j-1}^n] | \leq \\ \leq \frac{1}{2} [|\nu - \frac{1}{2}\mu| + \mu + (\nu + \frac{1}{2}\mu)] (\|v^n\|^2 + \|w^n\|^2). \end{aligned}$$

При этом получается условие устойчивости

$$\mu/2 + \max(\mu/2, \nu) < 1, \quad (11.28)$$

которое оказывается наилучшим из всех тех условий, которые практически можно получить с помощью энергетического мето-



Р и с. 11.3. Условия устойчивости для разностной схемы (11.22), получающиеся различными методами.

да ¹⁾. Как видно из рис. 11.3, оно иногда является менее ограничительным, чем даже условие (11.26). Так получается потому, что (11.26) не является наилучшим результатом, который

¹⁾ Указанная выше граница не может быть достигнута ни для каких сеточных функций v_j^n и w_j^n , если $\nu > \mu/2$. Но чтобы использовать это обстоятельство с целью ослабления неравенства (11.28), нужно рассмотреть приведение подобных членов в последней сумме. В действительности это снова приводит к методу Фурье, с помощью которого, как мы это знаем из теоремы Крайса о матрицах (см. гл. 4), может быть найден полный интервал устойчивости.

можно получить из (11.25): если мы положим $z = \sin^2(1/2 k \Delta x)$, то неравенство (11.25) перепишется в виде

$$\mu^2 z^2 + 4\nu^2 (z - z^2) \leq 1 \quad \text{при} \quad 0 \leq z \leq 1.$$

Максимум в левой части достигается при $z = 1$ для $2\nu^2 \leq \mu^2$ и равен в этом случае μ^2 ; в противном случае максимум равен $4\nu^4/(4\nu^2 - \mu^2)$. Следовательно, *необходимое и достаточное условие того, чтобы схема (11.22) была не только устойчивой, но и компоненты Фурье при этом не возрастали бы, имеет вид*

$$\begin{aligned} \mu < 1, \text{ если } \mu^2 \geq 2\nu^2, \\ (\mu/2\nu)^2 + \nu^2 \leq 1 \text{ в противном случае.} \end{aligned} \tag{11.29}$$

Таким образом, величина μ принимает либо все допустимые значения, либо интервал ее изменения увеличивается в 2ν раз по сравнению с тем, который допускается условием (11.26).

Глава 12

ОДНОМЕРНОЕ ДВИЖЕНИЕ ЖИДКОСТИ (ГАЗА)

§ 12.1. Введение

Конечноразностные методы расчета нестационарных течений большей частью основываются на уравнениях движения либо в форме Эйлера, либо в форме Лагранжа¹⁾. Хотя эти формы по существу эквивалентны, форма Лагранжа дает больше информации: она показывает, куда каждая частица жидкости ушла от первоначального положения. Это значительно облегчает расчет движения двух или более жидкостей с различными уравнениями состояния. Кроме того, все известные до недавнего времени устойчивые разностные схемы для уравнений в форме Эйлера были менее точными, чем схемы для уравнений в форме Лагранжа, которым и отдавалось предпочтение. Положение изменилось после того, как Лакс и Вендрофф предложили схему второго порядка точности, которая оказалась особенно эффективной для уравнений в форме Эйлера.

Течение жидкости часто характеризуется внутренними разрывами, такими, как например скачки, на которых требуется ставить специальные граничные условия (условия на скачке). Эти условия описываются известными уравнениями Ренкина — Гюгонио, но их применение на практике затруднено тем, что поверхности разрывов перемещаются по жидкости, и это движение не известно заранее, а определяется из самих дифференциальных уравнений и условий на скачке, так что численные расчеты проводятся для сложной системы неявных уравнений. Мы начнем с описания дифференциальных уравнений, которые справедливы для гладкой части течения, т. е. для задач без скачков или для течения между скачками. Методы расчета движения скачков будут описаны в § 12.9 и 12.11.

¹⁾ В форме Эйлера независимые пространственные переменные относятся к системе координат, фиксированной в пространстве, в котором движется жидкость, и течение характеризуется зависящим от времени полем скоростей, которое должно разыскиваться как решение краевой задачи. В форме Лагранжа независимые пространственные переменные относятся к системе координат, связанной с движущейся жидкостью, так что частицы жидкости перманентно отождествляются с их лагранжевыми координатами, тогда как их действительное положение в пространстве определяется зависимыми переменными, которые должны быть найдены. По поводу дальнейших деталей см. книгу Куранта и Фридрихса [1948] или работу Бете и др. [1944].

§ 12.2. Уравнения Эйлера

Пусть $\mathbf{X} = (X, Y, Z)$ — радиус-вектор точки в декартовой системе координат. Свойства жидкой среды характеризуются плотностью $\rho = \rho(\mathbf{X}, t)$, давлением $p = p(\mathbf{X}, t)$, внутренней энергией, отнесенной к единице массы, $\mathcal{E} = \mathcal{E}(\mathbf{X}, t)$, и скоростью жидкости $\mathbf{u} = \mathbf{u}(\mathbf{X}, t)$. Предположим, что уравнение состояния задано в виде

$$\mathcal{E} = F(p, \rho), \quad (12.1)$$

и предположим также, что силы вязкости, объемные силы (подобные тяготению), теплопроводность и источники энергии отсутствуют, хотя объемные силы и источники энергии можно было бы включить без существенных осложнений. В дальнейшем в связи со скачками мы введем некоторые фиктивные силы псевдовязкости и создадим возможность для учета теплопроводности, сохранив уравнение энергии в системе вместо того, чтобы исключить его, используя постоянства энтропии вдоль траекторий частиц жидкости.

Сохранение материи выражается следующим уравнением (иногда называемым уравнением «непрерывности» или «неразрывности»)¹⁾:

$$\left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \right) \rho = -\rho \nabla \cdot \mathbf{u}; \quad (12.2)$$

уравнением движения служит

$$\rho \left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \right) \mathbf{u} = -\nabla p, \quad (12.3)$$

а уравнением энергии является

$$\rho \left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \right) \mathcal{E} = -p \nabla \cdot \mathbf{u}. \quad (12.4)$$

В случае симметрии слоя, когда ничто не зависит от Y и Z , мы вместо уравнений (12.2) — (12.4) имеем систему

$$\begin{aligned} \left(\frac{\partial}{\partial t} + u \frac{\partial}{\partial X} \right) \rho &= -\rho \frac{\partial u}{\partial X}, \\ \rho \left(\frac{\partial}{\partial t} + u \frac{\partial}{\partial X} \right) u &= -\frac{\partial p}{\partial X}, \\ \rho \left(\frac{\partial}{\partial t} + u \frac{\partial}{\partial X} \right) \mathcal{E} &= -p \frac{\partial u}{\partial X} \end{aligned} \quad (12.5)$$

и уравнение состояния (12.1).

¹⁾ Вывод этих уравнений и общие рассуждения о гидродинамических процессах см. в книге Куранта и Фридрихса [1948].

В случае сферической симметрии вид уравнений мало меняется, и удобно, объединяя оба случая, записать уравнения в следующем виде:

$$\begin{aligned} \left(\frac{\partial}{\partial t} + u \frac{\partial}{\partial R} \right) \rho &= - \rho \frac{1}{R^{\alpha-1}} \frac{\partial}{\partial R} R^{\alpha-1} u, \\ \rho \left(\frac{\partial}{\partial t} + u \frac{\partial}{\partial R} \right) u &= - \frac{\partial p}{\partial R}, \\ \rho \left(\frac{\partial}{\partial t} + u \frac{\partial}{\partial R} \right) \mathcal{E} &= - \rho \frac{1}{R^{\alpha-1}} \frac{\partial}{\partial R} R^{\alpha-1} u, \end{aligned} \quad (12.6)$$

где при сферической симметрии $R = \sqrt{X^2 + Y^2 + Z^2}$, $\alpha = 3$, а при симметрии слоя $R = X$, $\alpha = 1$. В случае цилиндрической симметрии (когда поток инвариантен как относительно переносов, параллельных некоторой оси, так и относительно поворотов вокруг этой оси) получается $R = \sqrt{X^2 + Y^2}$, $\alpha = 2$.

§ 12.3. Разностные уравнения Эйлера

Конечноразностные уравнения можно строить различными способами; их выбор зависит от многих факторов, включая устойчивость, точность и отражение важных физических свойств, например законов сохранения. Метод Куранта, Изаксона и Риса [1952]¹⁾ был бы очень хорошим, но он не получил широкого распространения, по-видимому, вследствие необходимости приведения уравнений к нормальной или характеристической форме. В этом методе разности по времени берутся вперед, и каждое уравнение в нормальной форме аппроксимируется пространственными разностями вперед или назад согласно знаку наклона соответствующей характеристики. Устойчивость и сходимость в надлежащем образом ограниченной области могут быть доказаны в предположении, что $\Delta t / \Delta X$ нигде не превосходит наклона характеристик, что дает условие устойчивости

$$(|u| + c) \Delta t / \Delta X \leq 1,$$

где $c = c(X, t)$ — местная адиабатическая (точнее, изэнтропическая) скорость звука. Для изэнтропического процесса мы имеем $d\mathcal{E} + p d(1/\rho) = 0$ или, учитывая (12.1),

$$\frac{\partial F}{\partial p} dp + \frac{\partial F}{\partial \rho} d\rho - \frac{p}{\rho^2} d\rho = 0.$$

¹⁾ См. описание приложения этого метода к уравнению переноса в гл. 9, где он носит название метода «пространственных разностей вперед и назад».

Скорость звука c есть корень квадратный из производной $dp/d\rho$, вычисленной вдоль адиабаты; поэтому

$$c^2 = \frac{p/\rho^2 - \partial F/\partial \rho}{\partial F/\partial p}. \quad (12.7)$$

Метод Куранта, Изаксона и Риса может быть интерпретирован следующим образом. Вдоль каждой характеристики некоторая величина остается почти постоянной (в точности постоянной в изэнтропическом случае, являясь тогда соответствующим инвариантом Римана). Чтобы получить значение этой величины в точке сетки с индексами $n+1, j$, проводим из этой точки характеристику назад до пересечения с прямой $t = n\Delta t$ и затем получаем ее значение линейной интерполяцией между точками n, j и $n, j+1$ или между $n, j-1$ и n, j . Такого рода методы высокого порядка точности для полулинейных гиперболических систем построил Анзорге [1963] по аналогии с известным методом Адамса для обыкновенных дифференциальных уравнений. Многомерные обобщения были сделаны и изучены Бруном и Хааком [1958] и Тёрнигом [1963]. Аналогичный многомерный метод, использующий кохарактеристики, был предложен и изучен Зауэром [1963]. Эти многомерные обобщения будут кратко описаны в гл. 13.

Лакс и Келлер [1951] предложили схему, использующую подвижную сетку в плоскости X, t . Эта схема имеет только первый порядок точности, но оказалась полезной для обоснования существования и других свойств решений дифференциальных уравнений. Простой метод, подобный методу Лакса и Келлера, но одинаково пригодный как для гладких, так и для разрывных течений, был предложен Годуновым [1959]; он будет кратко описан в § 12.15.

Наиболее часто применяемые схемы (как для уравнений Лагранжа, так и для уравнений Эйлера) строятся по образцу ранее приведенной схемы (10.5) для волнового уравнения (первоначальная схема Куранта, Фридрихса и Леви). Эти схемы имеют то общее, что второе уравнение из системы (12.5) или системы (12.6) применяется для того, чтобы сначала по значениям u в момент t найти значения u в момент $t + \Delta t$ ¹⁾, а затем использовать полученные значения u в двух других уравнениях.

Для уравнений Эйлера такая схема окажется неустойчивой если не проявить осторожности в трактовке членов, содержа-

¹⁾ С точки зрения центрирования старые и новые значения u в действительности относились бы к моментам $t - \Delta t/2$ и $t + \Delta t/2$ соответственно, тогда как другие величины относились бы к моментам t и $t + \Delta t$, но мы предпочитаем сохранить наши постоянные обозначения.

щих $u \partial / \partial X$ ¹⁾. Неудовлетворительной, например, является аппроксимация $u \partial \rho / \partial X$ величиной

$$u_j^{n+1} \frac{\rho_{j+1}^n - \rho_{j-1}^n}{2 \Delta X},$$

и аналогично обстоит дело с другими членами такого типа. Это легко усмотреть, написав уравнения в вариациях для разностных уравнений и подсчитав их матрицу перехода в пределе для $\Delta t = \Delta X = 0$, т. е. матрицу $G(0, k)$. Одним из собственных значений этой матрицы является

$$1 + (u \Delta t / \Delta X) 2i \sin k \Delta X;$$

максимальная величина его модуля равна $(1 + 4(u \Delta t / \Delta X)^2)^{1/2}$, что превосходит 1 (если u не равно тождественно нулю), так что условие фон Неймана не может выполняться ни для какого фиксированного значения $\Delta t / \Delta X$. Способ Лелевье состоит в том, чтобы в таких членах применять пространственные разности вперед или назад при $u < 0$ и при $u > 0$ соответственно; так, приведенное выше выражение заменяется на

$$u_j^{n+1} \frac{\rho_{j+1}^n - \rho_j^n}{\Delta X}, \quad \text{если } u_j^{n+1} < 0,$$

$$u_j^{n+1} \frac{\rho_j^n - \rho_{j-1}^n}{\Delta X}, \quad \text{если } u_j^{n+1} \geq 0;$$

и аналогично аппроксимируются другие члены такого типа в уравнениях (12.5) или (12.6). Сначала из второго уравнения определяются u_j^{n+1} для всех j ; затем эти значения, которые лучше всего относить к моменту $t + \Delta t / 2$ ²⁾, используются в первом и третьем уравнениях для определения ρ и $\mathcal{E}$ в момент $t + \Delta t$ через их значения в момент t .

Окончательная схема (мы опускаем детали) устойчива, если $(|u| + c) \Delta t / \Delta X < 1$. Однако она имеет точность лишь первого порядка, а это недостаточно для большинства приложений. Разностные уравнения Лакса — Вендроффа, которым теперь отдается предпочтение, так как они имеют второй порядок точности, будут описаны в § 12.7.

§ 12.4. Уравнения Лагранжа

Пусть в случае симметрии слоя x означает координату X малого элемента жидкости в момент $t = 0$, и пусть $X = X(x, t)$ означает координату X того же элемента в момент t , так что

¹⁾ Лелевье, частное сообщение [1953].

²⁾ См. примечание ¹⁾ на стр. 290.

x и t рассматриваются как независимые переменные. Каждая частица жидкости отмечена некоторым значением x , называемым лагранжевой координатой частицы и не меняющимся при движении частицы. Очевидно, значение x , которым отмечается частица, может быть выбрано и многими другими способами, отличными от только что описанного. Например, в качестве x можно было бы взять значение координаты X в некоторый другой фиксированный момент $t = t_0$. Более того, если $g(x)$ — некоторая функция с непрерывной положительной производной, то переменную $x' = g(x)$ также можно взять в качестве такого значения; иначе говоря, если функция $h(x)$ непрерывна и положительна, то в качестве лагранжевой координаты вместо x можно взять величину

$$x' = \int h(x) dx.$$

Частный выбор, который является удобным и будет применяться впредь, таков:

$$x' = \frac{1}{\rho_0} \int \rho(x) dx,$$

где $\rho(x)$ — плотность в момент $t = 0$, а ρ_0 — постоянная, имеющая размерность плотности. Координата x' может быть истолкована как координата X в *конфигурации соотнесений* системы, в которой различные слои жидкости перемещаются и расширяются или сжимаются до тех пор, пока все они не будут иметь одну и ту же общую плотность ρ_0 . Эту конфигурацию системы мы будем называть *стандартной конфигурацией* соотнесений. С этого момента мы опустим штрих и будем писать x вместо x' . Аналогично, пусть для сферы r означает радиальную координату частицы жидкости в стандартной конфигурации соотнесений, в которой различные сферические слои жидкости движутся внутрь или наружу до тех пор, пока все они не будут иметь плотность ρ_0 , и пусть $R = R(r, t)$ — радиальная координата того же самого элемента жидкости в момент t в процессе движения. Величины r и R могут быть определены аналогично и в случае цилиндрической симметрии (под которой мы опять понимаем инвариантность как относительно переносов, параллельных некоторой оси, так и относительно поворотов вокруг этой оси).

Как и в случае эйлеровых координат, мы можем объединить формулы для случаев симметрии слоя, цилиндрической и сферической симметрии, вводя α и полагая $\alpha = 1, 2, 3$ соответственно, а также считая $r = x$ и $R = X$ в случае симметрии

слоя. Мы получаем три основных уравнения

$$\begin{aligned}\frac{\partial R}{\partial t} &= u, \\ \frac{\partial u}{\partial t} &= -V_0 \left(\frac{R(r, t)}{r} \right)^{\alpha-1} \frac{\partial p}{\partial r}, \\ \frac{\partial \mathcal{E}}{\partial t} &= -p \frac{\partial V}{\partial t},\end{aligned}\tag{12.8}$$

где $V_0 = 1/\rho_0$, и два вспомогательных уравнения

$$\begin{aligned}V &= V_0 \left(\frac{R(r, t)}{r} \right)^{\alpha-1} \frac{\partial R}{\partial r}, \\ \mathcal{E} &= F_1(p, V).\end{aligned}\tag{12.9}$$

Здесь V — удельный объем $1/\rho$, а u , p и $\mathcal{E}$ снова означают скорость, давление и внутреннюю энергию, приходящуюся на единицу массы, как и в случае эйлеровых координат, но теперь все они считаются функциями от r и t .

Отметим, что при исследовании устойчивости, которое будет проведено ниже, основными зависимыми переменными будут V , u и p (или $\mathcal{E}$); R является дополнительным переменным, при исключении производных которого с помощью первого уравнения (12.8) и первого уравнения (12.9) устанавливается связь между u и V . Чтобы убедиться в целесообразности такого выбора зависимых переменных, заметим, что для малых отклонений от состояния покоя (звуковых волн) в случае симметрии слоя ($\alpha = 1$) получается система

$$\frac{\partial R}{\partial t} = u, \quad \frac{\partial u}{\partial t} = (cV_0/V)^2 \frac{\partial^2 R}{\partial r^2},$$

которая по форме в точности совпадает с системой, приведенной в начале § 10.2 и оказавшейся непригодной; поэтому если R и u использовать для построения нормы в банаховом пространстве, как это было описано в § 10.2, то задача окажется некорректно поставленной. Практически это означает, что численные значения R должны находиться с большей относительной точностью (с большим числом верных значащих цифр), чем другие переменные, так как основной физический смысл имеют только временные и пространственные разности от R .

Иногда более удобно (так же как и в эйлеровых координатах) иметь уравнение состояния в форме, разрешенной относительно p , а именно $p = f(\mathcal{E}, V)$. Тогда это уравнение и первое уравнение (12.9) можно трактовать просто как обозначения; их

подстановка в (12.8) дает нам три дифференциальных уравнения в частных производных относительно трех зависимых переменных R , u и $\mathcal{E}$.

§ 12.5. Разностные уравнения Лагранжа

Как и в эйлеровых координатах, система является гиперболической, и конечноразностные уравнения могут быть взяты в форме, предложенной Курантом, Изаконом и Рисом [1952] или Лаксом и Келлером [1951], но на практике чаще всего берут за образец систему (10.5) для волнового уравнения или систему типа Лакса — Вендроффа. Несмотря на повышенную устойчивость, неявные схемы применяются редко, потому что для большинства задач гидродинамики система заметно изменяется за время $\Delta t = \Delta x/c$, и поэтому нет оснований для применения большего шага по времени. Однако они оказались полезными для некоторых астрофизических задач, в которых ожидаемое движение состоит из последовательности медленно меняющихся и близких к равновесию состояний, в задачах, где скорость звука в одной части системы много больше, чем в других частях (и поэтому состояние всегда близко к равновесию), и в приложениях к метеорологии, где скорость основных процессов существенно меньше скорости звука (см. Тёрнер и Вендрофф [1964]).

Обычная система имеет вид

$$\begin{aligned} \frac{R_j^{n+1} - R_j^n}{\Delta t} &= u_j^{n+1}, \\ \frac{u_j^{n+1} - u_j^n}{\Delta t} &= -V_0 \left(\frac{R_j^n}{r_j} \right)^{\alpha-1} \frac{p_{j+1/2}^n - p_{j-1/2}^n}{\Delta r} \\ \mathcal{E}_{j+1/2}^{n+1} - \mathcal{E}_{j+1/2}^n &= -p_{j+1/2}^n (V_{j+1/2}^{n+1} - V_{j+1/2}^n), \\ V_{j+1/2}^{n+1} &= V_0 \frac{(R_{j+1}^{n+1})^\alpha - (R_j^{n+1})^\alpha}{(r_{j+1})^\alpha - (r_j)^\alpha}, \\ p_{j+1/2}^{n+1} &= f(\mathcal{E}_{j+1/2}^{n+1}, V_{j+1/2}^{n+1}). \end{aligned} \tag{12.10}$$

Мы применяем r_j для обозначения j -го узла лагранжевой сетки. Если мы хотим сделать погрешность системы величиной порядка $O[(\Delta t)^2] + O[(\Delta r)^2]$, то третье уравнение должно быть центрировано путем замены $p_{j+1/2}^n$ некоторой лучшей аппроксимацией p при $t = t^{n+1/2}$; это может быть сделано различными способами, но за счет введения двух неизвестных $p_{j+1/2}^{n+1}$ и $\mathcal{E}_{j+1/2}^{n+1}$ в третье уравнение; при этом в общем случае требуется метод итераций для решения этого уравнения совместно с пятым.

Остальные уравнения системы (12.10) в действительности должны быть центрированы (хотя это не усматривается непосредственно), потому что u_j^{n+1} и u_j^n на самом деле связаны с моментами времени $t^{n+1/2}$ и $t^{n-1/2}$.

Типичными граничными условиями в точке $j = J$ являются:

на жесткой стенке $u_j^n = 0$ для всех n ;

на свободной поверхности $p_{j+1/2}^n = -p_{j-1/2}^n$ для всех n .

В силу второго из этих условий интерполированное значение p при $j = J$ обращается в нуль. Граничные условия можно применять также к средней точке интервала, например к точке $j = J + 1/2$, следующим образом:

на жесткой стенке $u_{j+1}^n = -u_j^n$ для всех n ;

на свободной поверхности $p_{j+1/2}^n = 0$ для всех n .

Поверхности раздела будут специально обсуждены ниже.

Система (12.10) оказывается явной, если неизвестные величины определяются с ее помощью в определенном порядке. Если все искомые величины известны при $t = t^n$ и если границы определяются значениями $j = 0$ и $j = J$, то этот порядок может быть таким:

величины:	уравнения:
u_j^{n+1} ($j = 1, 2, \dots, J - 1$)	второе из (12.10),
u_0^{n+1}, u_J^{n+1}	граничные условия,
R_j^{n+1} ($j = 0, 1, \dots, J$)	первое из (12.10),
$V_{j+1/2}^{n+1}$ ($j = 0, 1, \dots, J - 1$)	четвертое из (12.10),
$\mathcal{E}_{j+1/2}^{n+1}$ ($j = 0, 1, \dots, J - 1$)	третье из (12.10),
$p_{j+1/2}^{n+1}$ ($j = 0, 1, \dots, J - 1$)	пятое из (12.10).

Этот порядок, в том виде как он описан здесь, требовал бы пяти проходов по пространственной сетке (циклов); на практике они обычно объединяются в один цикл.

Чтобы исследовать устойчивость этих уравнений, сначала рассмотрим малые возмущения постоянного течения в случае симметрии слоя ($\alpha = 1$) и при этом заменим Δx на Δr . Образуя разность по времени для четвертого уравнения системы (12.10) и используя первое уравнение, можно исключить R и получить

$$V_{j+1/2}^{n+1} - V_{j+1/2}^n = V_0 \frac{\Delta t}{\Delta x} (u_{j+1}^{n+1} - u_j^{n+1}). \quad (12.11)$$

Линеаризуя третье уравнение системы (12.10) и используя соотношение

$$\Delta \mathcal{E} = \frac{\partial F_1}{\partial p} \Delta p + \frac{\partial F_1}{\partial V} \Delta V,$$

находим условие

$$\frac{\partial F_1}{\partial p} (p_{j+1/2}^{n+1} - p_{j+1/2}^n) + \left(p + \frac{\partial F_1}{\partial V} \right) (V_{j+1/2}^{n+1} - V_{j+1/2}^n) = 0,$$

которое связывает приращения функций p и V по времени. Так как основное течение постоянно, то то же самое условие связывает и их приращения по пространству; следовательно, второе из уравнений (12.10) дает

$$u_j^{n+1} - u_j^n = \frac{c^2}{V^2} V_0 \frac{\Delta t}{\Delta x} (V_{j+1/2}^n - V_{j-1/2}^n), \quad (12.12)$$

где c^2 определяется формулой

$$c^2 = \frac{V^2 (p + \partial F_1 / \partial V)}{\partial F_1 / \partial p},$$

эквивалентной выражению (12.7). (Уравнение состояния теперь записывается в виде (12.9), а не в виде (12.1)). Уравнения (12.11) и (12.12) образуют систему, имеющую тот же вид, что и система (10.5) для волнового уравнения; следовательно, условием устойчивости является неравенство

$$\frac{V_0}{V} \frac{c \Delta t}{\Delta x} \leq 1.$$

Исследование устойчивости для сферического и цилиндрического случаев несколько сложнее и приводит к условию

$$\frac{V_0}{V} \left(\frac{R}{r} \right)^{\alpha-1} \frac{c \Delta t}{\Delta r} \leq 1.$$

Его можно записать в более естественной форме, полагая

$$\Delta R = \frac{V}{V_0} \left(\frac{r}{R} \right)^{\alpha-1} \Delta r$$

и учитывая, что с точностью разностных формул ΔR является разностью значений $R(r, t)$ для смежных точек r_j и r_{j+1} и, следовательно, является действительным расстоянием между частицами жидкости, имеющими лагранжевы координаты r_j и r_{j+1} . Таким образом, в качестве условия устойчивости мы получаем соотношение

$$\frac{c \Delta t}{R_{j+1}^n - R_j^n} \leq 1 \quad \text{для всех } n, j. \quad (12.13)$$

При машинных вычислениях машина обычно автоматически проверяет выполнение неравенства (12.13) для всех (или для

избранных) значений n и j и останавливается в случае невыполнения этого неравенства. Конечно, скорость звука c также зависит от n и j и должна находиться из уравнения состояния; если уравнение состояния сложно, то можно взять приближенную формулу для c и — чтобы гарантировать устойчивость — заменить единицу в правой части неравенства несколько меньшим числом.

§ 12.6. Поверхности раздела в лагранжевых координатах

Если поверхность раздела двух сред совпадает с некоторым узлом сетки $j = J$, то уравнения (12.10) могут применяться таким же образом, как и ранее, только уравнения состояния по разные стороны от поверхности раздела будут различными. Однако это может вести к неточностям, и лучше дать специальный метод для поверхности раздела. Неточности возникают отчасти потому, что если плотности двух сред сильно различаются (как в случае жидкости и газа), то может оказаться необходимым брать совершенно различные по величине Δr по разные стороны поверхности раздела с соответственно плохим центрированием уравнений (12.10), а отчасти потому, что хотя градиент давления $\partial p / \partial r$ изменяется непрерывно при переходе через поверхность раздела (это является следствием принятого здесь определения лагранжевых переменных — см. ниже), *производная* градиента давления уже будет претерпевать разрыв, так что $\partial p / \partial r$ не может с точностью до второго порядка аппроксимироваться центральным разностным отношением. Мы выведем уточненную формулу для вычислений при наличии поверхности раздела.

Прежде всего, чтобы обеспечить непрерывность $\partial p / \partial r$, предположим, что начальные данные являются гладкими (обычно они аналитичны), за исключением, возможно, конечного числа точек, в которых выполняются различные группы условий, и что уравнение состояния является гладким. Тогда решение будет гладким везде, за исключением конечного числа кривых в плоскости x, t , представляющих собой скачки уплотнения, волны разрежения, слабые разрывы, поверхности раздела и т. п. (Мы не делаем попыток доказать эти утверждения или хотя бы разъяснить их точный смысл, но заметим, что аналогичные вопросы рассматривались Лаксом [1953] в связи с аналитичностью решений.) Мы утверждаем, что если исключить те моменты времени, в которые скачки уплотнения, волны разрежения и т. п. пересекают поверхности раздела, то $\partial p / \partial r$ остается непрерывной при переходе через поверхность раздела по следующим причинам. В малой окрестности поверхности

раздела решение непрерывно по обе ее стороны. Очевидно, R и u должны быть непрерывными по r при переходе через поверхность раздела (нарушение этого условия привело бы к образованию полостей или проникновению одной среды в другую). Их общие значения, скажем $R_0(t)$ и $u_0(t)$, дают положение и скорость самой поверхности раздела. Из гладкости по обе стороны, в которую мы включаем существование равномерных односторонних пределов функций и их производных с обеих сторон, следует, что $\partial u / \partial t$ с обеих сторон стремится к du_0/dt и, следовательно, является непрерывной. Из уравнения движения (второго из уравнений (12.8)) тогда непосредственно вытекает, что функция $\partial p / \partial r$ будет непрерывной.

Мы используем непрерывность $\partial p / \partial r$, чтобы получить приближенную формулу для ускорения поверхности раздела. Пусть поверхность раздела соответствует $j = J$, и пусть $\Delta_1 r$ и $\Delta_2 r$ — приращения r , применяемые слева и справа от $j = J$ соответственно. Давления в средних точках интервалов обозначим, как обычно, через $p_{j-3/2}^n$, $p_{j-1/2}^n$, $p_{j+1/2}^n$, $p_{j+3/2}^n$ и т. д., а через p_j^n обозначим давление на поверхности раздела (p конечно, непрерывно). Мы не знаем величины p_j^n , но она будет исключена из окончательной формулы. Слева от поверхности раздела мы имеем

$$\frac{p_{j-1/2}^n - p_{j-3/2}^n}{\Delta_1 r} \approx \frac{\partial p}{\partial r} \quad \text{при } r = r_{j-1},$$

$$\frac{p_j^n - p_{j-1/2}^n}{\Delta_1 r/2} \approx \frac{\partial p}{\partial r} \quad \text{при } r = r_{j-1/2},$$

и, следовательно, экстраполяция дает

$$\frac{4}{3} \frac{p_j^n - p_{j-1/2}^n}{\Delta_1 r/2} - \frac{1}{3} \frac{p_{j-1/2}^n - p_{j-3/2}^n}{\Delta_1 r} \approx \frac{\partial p}{\partial r} \quad \text{при } r = r_j.$$

Аналогично, справа от поверхности раздела получаем

$$\frac{4}{3} \frac{p_{j+1/2}^n - p_j^n}{\Delta_2 r/2} - \frac{1}{3} \frac{p_{j+3/2}^n - p_{j+1/2}^n}{\Delta_2 r} \approx \frac{\partial p}{\partial r} \quad \text{при } r = r_j.$$

Исключая p_j^n , находим

$$\left. \frac{\partial p}{\partial r} \right|_{r=r_j} \approx \frac{3(p_{j+1/2}^n - p_{j-1/2}^n) - 1/3(p_{j+3/2}^n - p_{j-3/2}^n)}{\Delta_1 r + \Delta_2 r}. \quad (12.14)$$

Это выражение нужно подставлять при $j = J$ вместо

$$\frac{p_{j+1/2}^n - p_{j-1/2}^n}{\Delta r}$$

в уравнение движения — второе из уравнений (12.10).

Формулу (12.14) можно применять даже в том случае, когда обе жидкости идентичны, т. е. когда нет реальной поверхности раздела, а при переходе через некоторую поверхность меняется лишь шаг Δr . Но в этом случае можно применять также и трехточечную формулу, включающую только $\rho_{j-1/2}^n$, $\rho_{j-1/2}^n$ и $\rho_{j+1/2}^n$ ($\rho_{j-1/2}^n$, $\rho_{j+1/2}^n$ и $\rho_{j+1/2}^n$), если принять во внимание непрерывность $\partial^2 \rho / \partial r^2$: нужно просто взять квадратичный трехчлен относительно r , принимающий заданные значения ρ в трех точках, и найти его производную при $r = r_j$. Трехточечная формула часто бывает более удобной с точки зрения программирования, но для реальных поверхностей раздела нужно пользоваться формулой (12.14).

§ 12.7 Дивергентная форма уравнений гидродинамики и уравнения Лакса — Вендроффа

Разностная схема, предложенная Лаксом и Вендроффом [1960], весьма удобна для решения уравнений Эйлера (12.6) и уравнений Лагранжа (12.8) в случае симметрии слоя. Как будет показано в следующей главе, вариант этой схемы, называемый *двухшаговой* схемой Лакса — Вендроффа, удобен для решения многомерных задач в форме Эйлера.

Сначала запишем уравнения гидродинамики в дивергентной форме, т. е. в виде законов сохранения. В форме Эйлера введем новые зависимые переменные

$$m = \rho u, \quad e = \rho (\mathcal{E} + 1/2 u^2); \quad (12.15)$$

ρ , m и e — масса, импульс и энергия на единицу объема соответственно. Тогда уравнения гидродинамики примут вид

$$\frac{\partial}{\partial t} \mathbf{U} + \frac{\partial}{\partial x} \mathbf{F}(\mathbf{U}) = 0, \quad (12.16)$$

где $\mathbf{U}$ и $\mathbf{F}(\mathbf{U})$ — векторы, определенные равенствами

$$\mathbf{U} = \begin{bmatrix} \rho \\ m \\ e \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} m \\ (m^2/\rho) + p \\ (e + p) m/\rho \end{bmatrix}. \quad (12.17)$$

Давление определяется уравнением состояния

$$p = f(\mathcal{E}, V) = f\left(\frac{e}{\rho} - \frac{m^2}{2\rho^2}, \frac{1}{\rho}\right), \quad (12.18)$$

где f — функция, введенная в § 12.4. В силу (12.18) $\mathbf{F}(\mathbf{U})$ действительно является функцией от $\mathbf{U}$. Система (12.16) называется *дивергентной формой уравнений гидродинамики*. Физическая интерпретация этого определения состоит в следующем:

каждая компонента $\mathbf{U}$ является величиной, отнесенной к единице объема; если слой заключен между x_1 и x_2 , то полное количество этой величины на единицу площади слоя равно интегралу от x_1 до x_2 от этой компоненты; но из (12.16) следует, что

$$\frac{d}{dt} \int_{x_1}^{x_2} \mathbf{U} dx = -\mathbf{F}(\mathbf{U})|_{x=x_2} + \mathbf{F}(\mathbf{U})|_{x=x_1}, \quad (12.19)$$

следовательно, соответствующая компонента $\mathbf{F}(\mathbf{U})|_x$ есть поток этой величины через единичную площадку в точке x в положительном направлении оси x . Законы сохранения массы, импульса и энергии снова потребуются нам в § 12.8, где из них будут получены условия Ренкина — Гюгонио на скачке.

В форме Лагранжа сначала исключается R из (12.8) и (12.9) (при этом предполагается симметрия слоя, т. е. $\alpha = 1$), что дает одно уравнение:

$$\frac{\partial V}{\partial t} = V_0 \frac{\partial u}{\partial x}. \quad (12.20)$$

Полную энергию на единицу массы, равную $\mathcal{E} + u^2/2$, обозначим через E . Тогда при

$$\mathbf{U} = \begin{bmatrix} V \\ u \\ E \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = V_0 \begin{bmatrix} -u \\ p \\ pu \end{bmatrix} \quad (12.21)$$

уравнения Лагранжа оказываются приведенными к дивергентной форме (12.16). Здесь $p = f(E - u^2/2, V)$, где f снова обозначает функцию, введенную в § 12.4. Как и в форме Эйлера, эти три уравнения выражают законы сохранения массы, импульса и энергии.

Уравнения Лакса — Вендроффа получаются с помощью разложения функции $\mathbf{U}$ в ряд Тейлора по t , т. е.

$$\mathbf{U}_j^{n+1} = \mathbf{U}_j^n + \Delta t \left(\frac{\partial \mathbf{U}}{\partial t} \right)_j^n + \frac{\Delta t^2}{2} \left(\frac{\partial^2 \mathbf{U}}{\partial t^2} \right)_j^n + \dots$$

Производные по t заменяются производными по x путем использования дифференциального уравнения (12.16) и получающегося из него уравнения

$$\frac{\partial^2 \mathbf{U}}{\partial t^2} = - \frac{\partial}{\partial t} \frac{\partial \mathbf{F}}{\partial x} = - \frac{\partial}{\partial x} \frac{\partial \mathbf{F}}{\partial t} = - \frac{\partial}{\partial x} \left(A \frac{\partial \mathbf{U}}{\partial t} \right) = \frac{\partial}{\partial x} \left(A \frac{\partial \mathbf{F}}{\partial x} \right),$$

где матрица $A = A(\mathbf{U})$ — матрица Якоби функции $\mathbf{F}(\mathbf{U})$, т. е. $A_{ij} = \partial F_i / \partial U_j$. Затем производные по x аппроксимируются

разностными отношениями, что приводит к системе уравнений Лакса — Вендроффа

$$U_j^{n+1} = U_j^n - \frac{1}{2} \frac{\Delta t}{\Delta x} (F_{j+1}^n - F_{j-1}^n) + \\ + \frac{1}{2} \left(\frac{\Delta t}{\Delta x} \right)^2 [A_{j+1/2}^n (F_{j+1}^n - F_j^n) - A_{j-1/2}^n (F_j^n - F_{j-1}^n)], \quad (12.22)$$

где через $A_{j+1/2}^n$ обозначена матрица $A(1/2 U_{j+1}^n + 1/2 U_j^n)$. Если матрица A постоянна, то эта система приводится к виду

$$U_j^{n+1} = U_j^n - 1/2 A \frac{\Delta t}{\Delta x} (U_{j+1}^n - U_{j-1}^n) + \\ + \frac{1}{2} \left(A \frac{\Delta t}{\Delta x} \right)^2 (U_{j+1}^n - 2U_j^n + U_{j-1}^n). \quad (12.23)$$

Мы укажем другую схему (двухшаговую схему Лакса — Вендроффа), которая также имеет второй порядок точности и сводится к (12.23) в случае постоянных коэффициентов. Она несколько проще для использования, потому что в нее не входит матрица A . Сначала в центрах прямоугольных ячеек в плоскости x, t вычисляются промежуточные значения

$$U_{j+1/2}^{n+1/2} = \frac{1}{2} (U_{j+1}^n + U_j^n) - \frac{\Delta t}{2 \Delta x} (F_{j+1}^n - F_j^n), \quad (12.24a)$$

где F_j^n обозначает $F(U_j^n)$ и т. д.; окончательные значения получаются затем из уравнения

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x} (F_{j+1/2}^{n+1/2} - F_{j-1/2}^{n+1/2}). \quad (12.24b)$$

Уравнение (12.24b) по своей структуре имеет второй порядок точности (остаточный член $O[(\Delta t)^3]$ при фиксированном отношении $\Delta t/\Delta x$); промежуточные значения получаются только с первым порядком точности, но разность в правой части (12.24b) имеет второй порядок точности, так как одинаковые члены взаимно уничтожаются.

Легко проверить, что в частном случае, когда $F(U) = AU$, подстановка (12.24a) в (12.24b) дает (12.23), как и утверждалось выше.

В случае постоянства коэффициентов условие устойчивости системы Лакса — Вендроффа легко устанавливается. Матрица перехода для системы (12.23) имеет вид

$$G = I - i \frac{\Delta t}{\Delta x} A \sin \alpha - \left(\frac{\Delta t}{\Delta x} A \right)^2 (1 - \cos \alpha), \quad (12.25)$$

где $\alpha = k\Delta x$. Матрица A имеет действительные собственные значения и полную совокупность линейно независимых собственных векторов, так как это основное свойство гиперболичности системы $\partial U/\partial t + A \partial U/\partial x = 0$, а уравнения гидродинамики

гиперболичны. Так как G — рациональная функция от матрицы A , она имеет те же самые собственные векторы, что и A ; следовательно, определитель Грама для этих собственных векторов не зависит от α и от $\Delta t/\Delta x$ и в силу условия 1 из § 4.11 условие фон Неймана не только необходимо, но и достаточно для устойчивости. Если λ — собственное значение матрицы A , то соответствующее собственное значение матрицы G равно

$$g = 1 - i \frac{\lambda \Delta t}{\Delta x} \sin \alpha - \left(\frac{\lambda \Delta t}{\Delta x} \right)^2 (1 - \cos \alpha).$$

Когда α меняется от 0 до 2π , g описывает эллипс в комплексной плоскости, и простой подсчет показывает, что этот эллипс лежит внутри единичной окружности при $|\lambda| \Delta t/\Delta x < 1$; для устойчивости это неравенство должно выполняться для всех собственных значений λ матрицы A . Заметим, что в точке $g = 1$ эллипс имеет касание второго порядка с единичной окружностью, т. е. они имеют в этой точке одинаковую кривизну. Это свидетельствует о втором порядке точности системы Лакса — Вендроффа и о недиссипативности системы дифференциальных уравнений, которая точно сохраняет амплитуду каждой компоненты Фурье при постоянной матрице A . С другой стороны, приведенная выше формула для g показывает, что $|g|^2 = 1 - 4\mu^2(1 - \mu^2)\sin^4(\alpha/2)$, где $\mu = |\lambda| \Delta t/\Delta x$, так что при $0 < \mu < 1$ порядок диссипативности разностной схемы в смысле Крайса равен четырем (см. § 5.4). В эйлеровой форме, которая будет рассмотрена вначале, матрица A , получаемая из (11.17), является довольно громоздкой, но ее собственные значения могут быть найдены косвенно. Действительно, исходные эйлеровы уравнения (12.5) также имеют вид $\partial U/\partial t + A \partial U/\partial x = 0$ с другой матрицей A , но эти две матрицы имеют одни и те же собственные значения, так как величины, обратные этим собственным значениям, являются тангенсами углов наклона характеристик в плоскости x, t . Чтобы найти эти собственные значения, подставим выражение (12.1) для $\mathcal{E}$ в третье уравнение системы (12.5) и исключим производные от p из этого уравнения с помощью первого уравнения (12.5); это дает

$$\frac{\partial p}{\partial t} + u \frac{\partial p}{\partial x} + \rho c^2 \frac{\partial u}{\partial x} = 0.$$

Тогда система принимает нужный вид $\partial U/\partial t + A \partial U/\partial x = 0$, где

$$A = \begin{bmatrix} u & p & 0 \\ 0 & u & 1/\rho \\ 0 & \rho c^2 & u \end{bmatrix}, \quad U = \begin{bmatrix} p \\ u \\ p \end{bmatrix}. \quad (12.26)$$

Собственные значения матрицы A равны u , $u + c$ и $u - c$; следовательно, условие устойчивости для уравнений Лакса — Вендроффа имеет вид

$$(|u| + c) \Delta t / \Delta X < 1, \quad (12.27)$$

как и для других разностных схем в форме Эйлера.

Аналогично в форме Лагранжа матрица A , получаемая непосредственно из (12.21), оказывается довольно громоздкой, однако третье уравнение можно переписать в виде

$$\frac{\partial p}{\partial t} + \frac{c}{V^2} V_0 \frac{\partial u}{\partial x};$$

тогда, взяв в качестве компонент U величины V , u и p , мы получаем матрицу

$$A = \begin{bmatrix} 0 & -V_0 & 0 \\ 0 & 0 & V_0 \\ 0 & V_0 (c/V)^2 & 0 \end{bmatrix}, \quad (12.28)$$

собственные значения которой равны 0 и $\pm c V_0 / V$, так что условием устойчивости является неравенство

$$\frac{V_0}{V} \frac{c \Delta t}{\Delta x} < 1, \quad (12.29)$$

как и для предыдущей схемы в форме Лагранжа.

Матрица (12.26) или (12.28) может быть приведена к диагональной форме преобразованием подобия PAP^{-1} (в качестве столбцов матрицы P^{-1} можно взять, например, нормированные собственные векторы матрицы A). Если систему дифференциальных уравнений, соответствующую (12.22), умножить на матрицу P , то она преобразуется к так называемой характеристической форме (см. Курант и Гильберт [1964] или Курант и Фридрихс [1950]). В переменных Эйлера уравнения в характеристической форме имеют вид

$$\begin{aligned} \left(\frac{\partial}{\partial t} + u \frac{\partial}{\partial X} \right) p - c^2 \left(\frac{\partial}{\partial t} + u \frac{\partial}{\partial X} \right) p &= 0, \\ \left(\frac{\partial}{\partial t} + (u + c) \frac{\partial}{\partial X} \right) p + pc \left(\frac{\partial}{\partial t} + (u + c) \frac{\partial}{\partial X} \right) u &= 0, \\ \left(\frac{\partial}{\partial t} + (u - c) \frac{\partial}{\partial X} \right) p - pc \left(\frac{\partial}{\partial t} + (u - c) \frac{\partial}{\partial X} \right) u &= 0, \end{aligned} \quad (12.30)$$

а в форме Лагранжа они запишутся как

$$\begin{aligned} \frac{\partial p}{\partial t} + \frac{c^2}{V^2} \frac{\partial V}{\partial t} &= 0, \\ \left(\frac{\partial}{\partial t} + c \frac{V_0}{V} \frac{\partial}{\partial x} \right) p + \frac{c}{V} \left(\frac{\partial}{\partial t} + c \frac{V_0}{V} \frac{\partial}{\partial x} \right) u &= 0, \\ \left(\frac{\partial}{\partial t} - c \frac{V_0}{V} \frac{\partial}{\partial x} \right) p - \frac{c}{V} \left(\frac{\partial}{\partial t} - c \frac{V_0}{V} \frac{\partial}{\partial x} \right) u &= 0, \end{aligned} \quad (12.31)$$

Достоинство двухшаговой процедуры состоит в том, что разностные уравнения (12.24а) и (12.24б) тоже записаны в виде законов сохранения, так что величины, которые должны сохраняться согласно дифференциальным уравнениям (12.16), в точности сохраняются и в разностной форме. Действительно, вместо интеграла в левой части (12.19) мы имеем сумму

$$S^n = \sum_{j=j_1}^{j_2} U_j^n \Delta x,$$

а вместо соотношения (12.19) — равенство

$$\frac{S^{n+1} - S^n}{\Delta t} = -F_{j_2+1/2}^{n+1/2} + F_{j_1-1/2}^{n+1/2}, \quad (12.32)$$

которое непосредственно следует из суммирования по j уравнений (12.24б).

Еще одно достоинство этого метода заключается в том, что матрица A и ее производные не входят в явном виде (это упрощает вычисления), что особенно важно для многомерных задач (см. гл. 13). Вычислительная программа должна содержать подпрограмму для вычисления компонент вектора $F(U)$ по компонентам вектора U , которая в общем случае весьма проста.

Класс разностных схем для уравнения (12.16), получающихся как последовательные приближения к неявному уравнению

$$U_j^{n+1} - U_j^n + \frac{\Delta t}{4\Delta x} (F_{j+1}^{n+1} + F_{j+1}^n - F_{j-1}^{n+1} - F_{j-1}^n) = 0,$$

был рассмотрен в работе Гари [1962b].

§ 12.8. Условия на скачке

Известно (см., например, Курант и Фридрихс [1950]), что движение идеальной жидкости часто характеризуется кривыми в плоскости x, t (или в более общем случае поверхностями в пространстве X, t), на которых некоторые из зависимых величин претерпевают разрыв, но имеют односторонние предельные значения с обеих сторон. (Даже если течение сначала является гладким, разрывы могут развиваться впоследствии за конечное время, по истечении которого уже не будет существовать гладкого однозначного решения системы (12.8).) Примерами служат поверхность раздела, слабый разрыв (где ρ и $\mathcal{E}$ разрывны, а p и u непрерывны), скачок уплотнения (где $\rho, \mathcal{E}, p, u$ разрывны) и фронт волны разрежения (где разрывны лишь некоторые производные). На таких разрывах дифференциальные уравнения (лишающиеся смысла) должны быть заменены условиями на скачке, которые служат внутренними гра-

ничными условиями и нужны для обеспечения единственности решения.

В случае поверхности раздела, слабого разрыва или фронта волны разрежения нужны лишь условия непрерывности некоторых переменных и не требуется изменения конечноразностной системы, хотя, как уже было указано, для поверхности раздела может быть желательной модификация разностной системы с целью повышения точности. Но если имеет место скачок, то немодифицированные разностные уравнения (12.10), за исключением некоторых предельных случаев, не дают даже приближенно правильных результатов. Эти модифицированные уравнения изучались фон Нейманом при экспериментальных вычислениях во время войны ¹⁾.

Мы установим условия на скачке в случае симметрии слоя, обозначив через $x = \xi(t)$ лагранжеву координату скачка. Масса жидкости, проходящая через единицу поверхности скачка за единицу времени, равна

$$M = \rho_0 d\xi/dt. \quad (12.33)$$

Пусть $V = 1/\rho$ означает удельный объем (условия на скачке обычно удобнее выражать через V , а не через ρ), и пусть индексы 1 и 2 определяют предельные значения при $x \rightarrow \xi$ слева и справа, так что $u_1, V_1, \rho_1, \mathcal{E}_1$ описывают состояние непосредственно за скачком, а $u_2, V_2, \rho_2, \mathcal{E}_2$ — непосредственно перед ним, если скачок движется в положительном направлении оси x и, наоборот, если скачок распространяется в отрицательном направлении оси x . Закон сохранения массы соответствует непрерывности $X(x, t)$ на скачке, так что производная от X по времени вдоль кривой $x = \xi(t)$, а именно

$$\frac{d\xi}{dt} \frac{\partial X}{\partial x} + \frac{\partial X}{\partial t},$$

должна быть одинаковой по обе стороны от скачка, но $\partial X/\partial x = \rho_0 V$, а $\partial X/\partial t = u$, так что

$$MV_1 + u_1 = MV_2 + u_2. \quad (12.34)$$

Закон сохранения количества движения гласит, что приращение количества движения жидкости, проходящей за единицу

¹⁾ У нас не должно создаться впечатление, что фон Нейман нахеревался таким путем получить правильный результат. Он указал, что немодифицированные уравнения недопустимы ввиду более сильного по сравнению с действительным возрастания энтропии при переходе через скачок, и исследовал ошибку, возникающую при игнорировании этого факта. Тем не менее полное игнорирование скачков при вычислениях некоторые исследователи называют «методом фон Неймана для расчета скачков». Как показано в § 12.14, уравнения Лакса — Вендроффа иногда можно использовать в немодифицированной форме при переходе через скачок, потому что в них заложен внутренний диссипативный механизм, обеспечивающий необходимое возрастание энтропии на скачке.

времени через единицу поверхности скачка, равно разности давлений, т. е.

$$M(u_1 - u_2) = p_1 - p_2. \quad (12.35)$$

Закон сохранения энергии утверждает, что приращение энергии этой жидкости равно результирующей работе сил давления за единицу времени, т. е.

$$M\left(\mathcal{E}_1 + \frac{u_1^2}{2} - \mathcal{E}_2 - \frac{u_2^2}{2}\right) = p_1 u_1 - p_2 u_2. \quad (12.36)$$

Уравнения (12.33) — (12.36) можно комбинировать различными способами. В результате одной из таких комбинаций получаются следующие уравнения:

$$\begin{aligned} \rho_0 \frac{d\xi}{dt} &= \sqrt{\frac{p_1 - p_2}{V_2 - V_1}}, \\ u_1 - u_2 &= \sqrt{(p_1 - p_2)(V_2 - V_1)}, \\ \mathcal{E}_1 - \mathcal{E}_2 &= \frac{p_1 + p_2}{2} (V_2 - V_1), \end{aligned} \quad (12.37)$$

называемые уравнениями Ренкина — Гюгонио или просто уравнениями Гюгонио¹⁾. Они сохраняют силу также для сферического и цилиндрического скачков, если $r = \xi(t)$ является лагранжевой координатой фронта скачка. Более того, они остаются локально справедливыми и в общем случае гладкой криволинейной поверхности скачка, если $\rho_0 d\xi/dt$ означает массу жидкости, проходящей через единицу поверхности скачка за единицу времени, а u_1 и u_2 означают нормальные по отношению к поверхности скачка составляющие скорости; тангенциальные составляющие скорости при переходе через скачок не претерпевают разрыва.

§ 12.9. «Подгонка» скачка²⁾

Уравнение Гюгонио совместно с условиями непрерывности $X(x, t)$ или $R(r, t)$ при переходе через скачок позволяют установить граничные условия для дифференциальных уравнений по обе стороны от скачка, а также определить движение самого скачка. Они использовались непосредственно при численных расчетах (неопубликованная работа, Лос-Аламос, а также работы Томаса [1954], Гари [1962а], Кворлса [1964]). Мы наметим после-

¹⁾ Вывод уравнений (12.37) из предыдущих уравнений содержит известную неопределенность в выборе знака перед квадратным корнем. Нужно брать знак плюс, как это сделано у нас, если скачок движется вправо, и знак минус, если он движется влево.

²⁾ В оригинале «shock fitting». — Прим. ред.

довательность вычислений для частного случая, когда среда перед скачком находится в состоянии покоя и в равновесии; тогда u_2 , V_2 , p_2 , $\mathcal{E}_2$ — известные постоянные ($u_2 = 0$), в то время как u_1 , V_1 , p_1 , $\mathcal{E}_1$ — функции от t , которые должны быть определены в результате расчета. Для каждого n должны быть вычислены положение скачка $\xi^n = \xi(t^n)$ (оно в общем случае не совпадает ни с одной из точек сетки) и значения величин u_1 , V_1 , p_1 , $\mathcal{E}_1$ на скачке, т. е. в точке $\xi = \xi^n$ при $t = t^n$; мы обозначим их через u_1^n , V_1^n , p_1^n , $\mathcal{E}_1^n$ и просим читателя не смешивать их с u_j^n при $j = 1$ и т. д.

Если все величины известны в момент $t = t^n$, то уравнения Ренкина — Гюгонио (12.37), в которых каждая величина снабжена верхним индексом $n + 1$, совместно с уравнением вида

$$\xi^{n+1} - \xi^n = (\Delta t/2) (\dot{\xi}^n + \dot{\xi}^{n+1}), \quad (12.38)$$

где $\dot{\xi} = d\xi/dt$, и с уравнением состояния

$$p_1^{n+1} = f(\mathcal{E}_1^{n+1}, V_1^{n+1}) \quad (12.39)$$

образуют систему пяти уравнений с шестью неизвестными ξ^{n+1} , $\dot{\xi}^{n+1}$, V_1^{n+1} , $\mathcal{E}_1^{n+1}$, p_1^{n+1} , u_1^{n+1} . Необходимо дополнить ее еще одним уравнением, чтобы отразить влияние на скачок движения среды непосредственно за скачком; для этой цели в разностной форме записывается одно из основных дифференциальных уравнений или некоторая их комбинация, причем пространственные разности берутся между значениями в точке $x = \xi^{n+1}$ и значениями в ближайших точках сетки позади скачка, в временные разности берутся между значениями в точках ξ^{n+1} , t^{n+1} и ξ^n , t^n , чтобы описать перемещение скачка между моментами t^n и t^{n+1} .

Для того чтобы решить, каким дифференциальным уравнением или какой их комбинацией дополнить таким образом условия на скачке, воспользуемся характеристической формой этих уравнений, приведенной в конце § 12.7¹⁾. Кривые в плоскости x , t , для которых величины, обратные к тангенсам углов наклона, являются собственными значениями матрицы A , называются *характеристиками*; через каждую точку проходит три характеристики, одна из которых соответствует распространению звуковых импульсов вверх по течению («опережающая

¹⁾ Сначала было эмпирически установлено, что может быть использовано уравнение непрерывности, и вполне возможно, что этот выбор является наилучшим. Проведенный в основном тексте анализ направлен на то, чтобы пролить некоторый свет на результаты возможного выбора, учитывая прежде всего последующие обобщения на значительно более трудные многомерные задачи (см. гл. 13).

характеристика»), другая — их распространению вниз по течению («запаздывающая характеристика»), а третья, лежащая между ними, — истинному движению частицы.

В известном смысле характеристики являются носителями информации. Например, в силу третьего из уравнений (12.8) энтропия S остается постоянной вдоль средней характеристики, так что знание S в какой-либо точке дает информацию об этой величине в любой другой точке, лежащей на этой характеристике. Кроме того, если энтропия S постоянна для всего потока, так что V и c могут быть выражены как функции от p , то уравнения (12.31) показывают, что величины $\int (V/c) dp \pm u$ (они называются *инвариантами Римана*) остаются постоянными вдоль опережающей и запаздывающей характеристик соответственно. Даже если S не является постоянной, за исключением траекторий движения частиц, для небольших промежутков времени величины, аналогичные инвариантам Римана, остаются почти постоянными вдоль опережающей и запаздывающей характеристик: например, одна из возможных разностных аппроксимаций второго из уравнений (12.31) гласит, что вдоль опережающей характеристики величина $p^{n+1} + (c/V)^{n+1/2} u^{n+1}$ приближенно равна $p^n + (c/V)^{n+1/2} u^n$.

При «подгонке» скачка неполная информация, которую дают условия на скачке, должна быть дополнена добавочной информацией, приходящей на скачок в момент t^{n+1} от находящейся за ним среды. Только одна из трех характеристик в состоянии сделать это, а именно та, которая представляет опережающий звуковой сигнал, догоняющий скачок во времени; если следовать назад из точки скачка, то две другие характеристики пройдут перед положением скачка в предшествующий момент времени, так что, например, средняя характеристика дает нам значение энтропии непосредственно перед скачком, а не интересующее нас значение непосредственно за скачком. Поэтому в случае скачка, движущегося вправо, для дополнения условий на скачке пригодны только второе из уравнений (12.31) в разностной форме. (Конечно, вопрос о том, не будет ли некоторая другая комбинация исходных уравнений в нехарактеристической форме столь же удобной или даже лучшей, остается открытым.) Получившиеся уравнения, которые нужно решить для нахождения шести неизвестных в точке ξ^{n+1} , t^{n+1} , нелинейны и довольно сложны, так что для их решения нужно использовать итерационный метод. Уравнения должны быть модифицированы, когда скачок проходит через точку сетки, т. е. когда ξ^n и ξ^{n+1} принадлежат различным интервалам пространственной сетки.

Из этого наброска схемы расчета ясно, вероятно, что «подгонка» скачка неудобна из-за ее сложности. Было установлено, что она применима в случае, когда скачок распространяется

в ранее невозмущенной среде, т. е. когда u_2 , p_2 , V_2 и $\mathcal{E}_2$ заранее известны, и что в общем случае его применение весьма затруднительно. Для более сложных задач возникают дополнительные трудности, потому что скачки могут спонтанно возникать внутри среды (см., например, Калкин [1942] или Курант и Фридрихс [1950]), и не существует метода, который позволил бы определить, где и когда можно использовать «подгонку» скачка, за исключением, возможно, постоянной проверки результатов расчета достаточно опытным специалистом.

Чтобы обойти эти трудности, фон Нейман и Рихтмайер [1950] разработали приближенный метод решения задач гидродинамики (излагаемый в трех следующих параграфах), в котором скачки учитываются автоматически, где бы и когда бы они ни возникали. Этот метод основывается на некоторых физических понятиях, относящихся к реальной (и поэтому не идеальной) жидкости, а не на непосредственном применении условий Гюгонио на скачке при вычислениях. В частности, явления диссипации, подобные вязкости и теплопроводности, оказывают на скачок сглаживающее влияние, так что поверхность разрыва заменяется тонким переходным слоем, в котором величины меняются быстро, но без разрыва. Дифференциальные уравнения, учитывающие диссипацию, применимы как в этом слое, так и вне него, так что в граничных условиях на скачке нет необходимости. В то же время основные законы сохранения, на которых базируются уравнения Гюгонио, остаются в силе, а условия на скачке выполняются для переходного слоя, толщина которого считается малой по сравнению с другими размерами, фигурирующими в задаче. Аналогичный метод был рассмотрен в работе Ладфорда, Полячека и Зегера [1953]. Более совершенный метод был предложен Лаксом и Вендроффом [1959]; их результаты будут изложены в § 12.14.

§ 12.10. Влияние диссипации

Влияние диссипации на скачок изучалось различными исследователями, включая Рэля, Беккера [1922] и Томаса [1944]. Беккер изучал влияние теплопроводности и вязкости. Он показал, что при наличии теплопроводности температура претерпевает гладкое изменение при переходе через скачок и что для скачков интенсивности¹⁾ менее критической давление и плотность также изменяются гладко, в то время как для более сильных скачков переход давлений и плотностей от начальных к конечным значениям осуществляется частично за счет гладких переходов и частично за счет разрывов. Однако если

¹⁾ Удобной мерой интенсивности скачка является отношение конечного давления к начальному.

допускается наличие вязкости, то при любой интенсивности скачка все величины меняются гладко в области скачка. В любом случае толщина переходного слоя пропорциональна коэффициенту диссипации, так что при стремлении к нулю теплопроводности и вязкости характер изменения величин приближается к разрывному в соответствии с теорией Гюгонио.

Идея фон Неймана и Рихтмайера состояла во введении чисто искусственной диссипации такого вида и такой величины, что скачок заменяется гладким переходом на протяжении небольшого числа (скажем, трех или четырех) интервалов Δr пространственного переменного, и в построении системы конечноразностных уравнений, учитывающих диссипацию и позволяющих обойтись без «подгонки» скачков. При вычислениях по таким уравнениям скачки выявляются автоматически как слои быстрого изменения величин, при переходе через которые величины претерпевают изменения, близкие к скачкам при переходе через фронт разрыва, причем эти слои движутся в жидкости со скоростями, близкими к истинным скоростям скачков уплотнения.

Очевидно, что механизм типа вязкости более удобен для наших целей, чем механизм типа теплопроводности, потому что при наличии вязкости все величины меняются гладко. Однако Беккер показал, что в случае обыкновенной вязкости, при которой напряжение пропорционально скорости сдвига и которая поэтому представляется в уравнениях линейными членами, толщина переходного слоя зависит от интенсивности скачка, стремясь к нулю для очень сильных скачков и к бесконечности для очень слабых скачков. Но мы хотим, чтобы для всех скачков толщина переходного слоя была примерно одинаковой, а именно составляла около $3\Delta r$ или $4\Delta r$, и поэтому вводим *квадратичные* члены в дифференциальное уравнение; это эквивалентно применению малого коэффициента вязкости для слабых скачков и большого коэффициента вязкости для сильных скачков. Ниже будет показано, что таким путем получается толщина, не зависящая от интенсивности скачка¹⁾. Другие

¹⁾ Возможно, фон Нейману и Рихтмайеру не стоило заниматься этим вопросом. Метод Лакса [1954] дает тот же эффект путем введения линейного члена, соответствующего вязкости. Действительно, Лакс нашел, что толщина переходного слоя меняется с интенсивностью скачка, но не настолько, чтобы препятствовать практическим вычислениям. Однако метод фон Неймана и Рихтмайера имеет то преимущество, что члены, соответствующие *псевдовязкости*, будучи квадратичными относительно скоростей деформаций, являются в действительности очень малыми в гладкой части течения между скачками, где желательно аппроксимировать поведение идеальной жидкости. Это желательно потому, что порядок величины действительной вязкости реальной жидкости меньше порядка величины вязкости, необходимой для образования скачков нужной толщины.

свойства этих дополнительных членов обеспечивают независимость толщины переходного слоя также от давления и плотности той среды, в которой распространяются скачки.

Уравнения, содержащие члены с такой псевдовязкостью, сначала будут записаны в дифференциальной форме для случая симметрии слоя. В лагранжевых переменных эти уравнения имеют вид

$$\begin{aligned}\frac{\partial X}{\partial t} &= u, \\ \frac{\partial u}{\partial t} &= -\frac{1}{\rho_0} \frac{\partial (p + q)}{\partial x}, \\ \frac{\partial \mathcal{E}}{\partial t} &= -(p + q) \frac{\partial V}{\partial t}, \\ V &= \frac{1}{\rho_0} \frac{\partial X}{\partial x}, \\ \mathcal{E} &= F(p, \rho).\end{aligned}\tag{12.40}$$

Здесь $V = 1/\rho$ — удельный объем, а q — псевдовязкое давление, задаваемое соотношением¹⁾

$$q = \begin{cases} (l^2/V) (\partial u / \partial x)^2 & \text{при } \partial u / \partial x < 0, \\ 0 & \text{при } \partial u / \partial x \geq 0, \end{cases}\tag{12.41}$$

где постоянная l имеет размерность длины.

Для того чтобы продемонстрировать влияние дополнительных членов, исследуем структуру плоского стационарного скачка и будем искать решения дифференциальных уравнений, в которых p , q , u , $\mathcal{E}$ и V зависят от x и t только в комбинации $\omega = x - st$, где s — скорость скачка в лагранжевых координатах, т. е. исследуем решения, имеющие вид распространяющейся волны. Кроме того, мы возьмем уравнение состояния в простейшем виде, соответствующем идеальному газу, $\mathcal{E} = pV/(\gamma - 1)$. В рассматриваемом случае симметрия слоя первое и четвертое уравнения (12.40) можно скомбинировать в одно уравнение

$$\frac{\partial V}{\partial t} = \frac{1}{\rho_0} \frac{\partial u}{\partial x},$$

¹⁾ Это выражение для q включает предложение М. Розенблюта (неопубликованное сообщение, Лос-Аламос), чтобы q обращалось в нуль, когда жидкость расширяется. Мы полагали сначала

$$q = -\frac{l^2}{V} \left| \frac{\partial u}{\partial x} \right| \frac{\partial u}{\partial x},$$

что соответствует формуле (12.41) при сжатии и дает отрицательные значения q при расширении. Для большинства задач обе формулы удовлетворительны, потому что величина q чрезвычайно мала (за исключением области скачка), но если имеет место сильное расширение (как в случае свободной поверхности), то формула Розенблюта лучше.

из которого X исключено. Подставляя в него и в другие уравнения решение в виде распространяющейся волны, находим

$$\begin{aligned}\rho_0 s \frac{dV}{d\omega} + \frac{du}{d\omega} &= 0, \\ \rho_0 s \frac{du}{d\omega} - \frac{d(p+q)}{d\omega} &= 0, \\ \frac{d\mathcal{E}}{d\omega} + (p+q) \frac{dV}{d\omega} &= 0.\end{aligned}$$

Пусть $M = \rho_0 s$ означает массу жидкости, проходящей через единицу поверхности за единицу времени; тогда решение этих уравнений можно записать в виде

$$\begin{aligned}MV + u &= C_1, \\ M^2V + p + q &= C_2, \\ \frac{1}{2}M^2V^2 + \mathcal{E} + (p+q)V &= C_3,\end{aligned}\tag{12.42}$$

где C_1 , C_2 и C_3 — постоянные интегрирования.

Если теперь мы рассмотрим решение¹⁾, для которого u , V , p , q асимптотически приближаются к некоторым значениям (обозначенным соответственно индексами 1 и 2) при $x \rightarrow -\infty$, и $+\infty$ и для которого производные этих величин стремятся к нулю при $x \rightarrow \pm\infty$, так что в частности $q \rightarrow 0$, то из уравнений (12.42) следуют равенства

$$\begin{aligned}M(V_1 - V_2) + u_1 - u_2 &= 0, \\ M^2(V_1 - V_2) + p_1 - p_2 &= 0, \\ \frac{1}{2}M^2(V_1^2 - V_2^2) + \mathcal{E}_1 - \mathcal{E}_2 + p_1V_1 - p_2V_2 &= 0,\end{aligned}$$

из которых после небольших преобразований получаются уравнения Гюгонио (12.37).

Из второго и третьего уравнений (12.42) и уравнения состояния получаем

$$\mathcal{E} = \frac{1}{2}M^2V^2 - C_2V + C_3 = pV/(\gamma - 1),$$

откуда с учетом второго уравнения (12.42) имеем

$$qV = -\frac{1}{2}(\gamma + 1)M^2V^2 + C_4V + C_5,$$

где C_4 и C_5 — некоторые постоянные. Но так как $q = 0$ при $V = V_1$ и при $V = V_2$, квадратичный трехчлен относительно V в правой части последнего равенства определяется однозначно, а именно

$$qV = \frac{1}{2}(\gamma + 1)M^2(V_2 - V)(V - V_1).\tag{12.43}$$

¹⁾ Общее решение такого рода будет приведено ниже.

Мы можем теперь заключить, что для решения в виде распространяющейся волны всегда будет $du/dw \leq 0$. Действительно, в интервале, где $du/dw \geq 0$, мы имеем $q = 0$ в силу (12.41), а отсюда следует, согласно (12.43), что в таком интервале V имеет либо постоянное значение V_2 , либо постоянное значение V_1 . Но из $V = \text{const}$ следует, что $u = \text{const}$, так как $\rho_0 \partial V / \partial t = du/dx$, поэтому если $du/dw \geq 0$, то $du/dw = 0$. *Распространяющаяся волна соответствует сжатию жидкости и никогда не соответствует ее расширению.* Уравнения Гюгонио (12.37) допускают отрицательные скачки (внезапные расширения) наряду с положительными, тогда как другие совершенно ясные физические соображения¹⁾ запрещают их существование, и приятно видеть, что уравнения, содержащие члены с псевдовязкостью, допускают только положительные скачки.

Так как $du/dw \leq 0$, условие (12.41) дает для распространяющейся волны

$$q = \frac{l^2}{V} \left(\frac{du}{dw} \right)^2 = \frac{l^2 M^2}{V} \left(\frac{dV}{dw} \right)^2,$$

и поэтому мы получаем уравнение

$$\left(l \frac{dV}{dw} \right)^2 = \frac{1}{2} (\gamma + 1) (V_2 - V) (V - V_1)$$

с общим решением

$$V = \frac{V_2 + V_1}{2} + \frac{V_2 - V_1}{2} \sin \sqrt{(\gamma + 1)/2} (\omega - \omega_0)/l,$$

где ω_0 постоянно. Но так как $du/dw \leq 0$ или $dV/dw \geq 0$, мы можем использовать только половину этой синусоидальной волны, которая должна непрерывно сопрягаться с двумя другими решениями $V \equiv V_1$ и $V \equiv V_2$. Выбирая для удобства начало отсчета так, чтобы $\omega_0 = 0$, мы получаем в конце концов решение в виде распространяющейся волны

$$V = \begin{cases} \frac{V_2 + V_1}{2} + \frac{V_2 - V_1}{2} \sin \sqrt{(\gamma + 1)/2} (x - st)/l \\ \quad \text{при } |x - st| \leq \pi \sqrt{2l}/(2 \sqrt{\gamma + 1}), \\ V_1 \quad \text{при } x - st \leq -\pi \sqrt{2l}/(2 \sqrt{\gamma + 1}), \\ V_2 \quad \text{при } x - st \geq +\pi \sqrt{2l}/(2 \sqrt{\gamma + 1}). \end{cases} \quad (12.44)$$

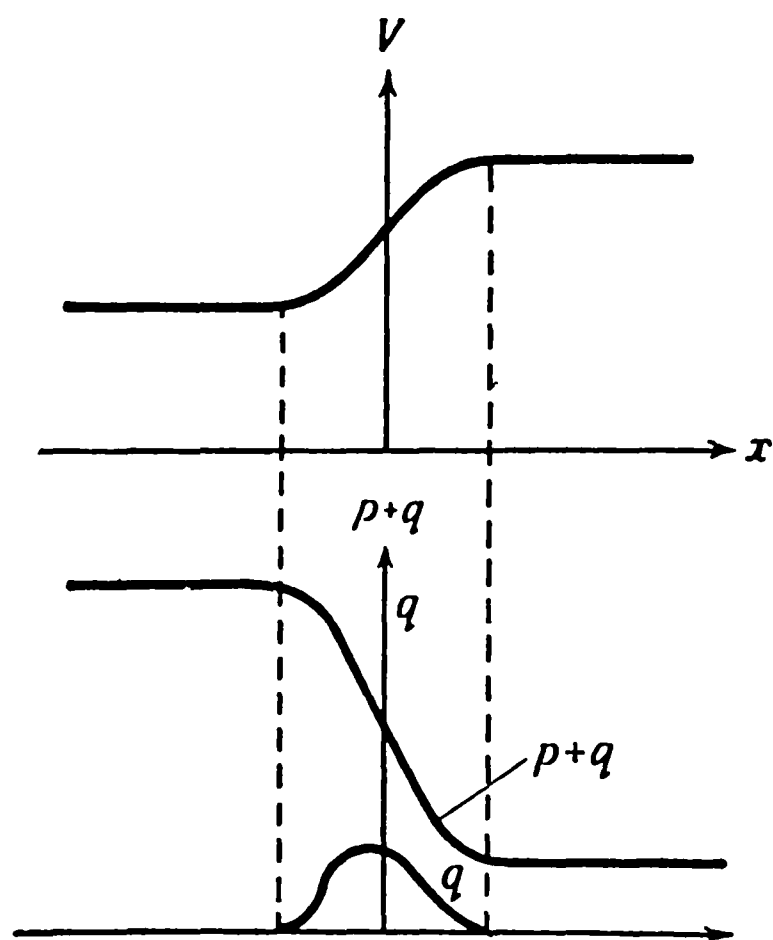
¹⁾ Отрицательный скачок мог бы явиться результатом убывания энтропии без соответствующего возрастания ее где-либо в другом месте, что противоречит второму закону термодинамики. Соображения устойчивости (фон Неймана) показывают, что отрицательный скачок сразу же распадается в веер разрежения, в котором изменения происходят непрерывно и энтропия не меняется.

Аналогично, полное и псевдовязкое давления в области скачка, т. е. при $|x - st| \leq \pi \sqrt{2l}/(2 \sqrt{\gamma + 1})$, даются формулами

$$p + q = \frac{p_1 + p_2}{2} - \frac{p_1 - p_2}{2} \sin \sqrt{(\gamma + 1)/2} (x - st)/l,$$

$$q = \frac{p_1 - p_2}{2} \frac{\gamma + 1}{2} \frac{[\cos \sqrt{(\gamma + 1)/2} (x - st)/l]^2}{(V_2 + V_1)/(V_2 - V_1) + \sin \sqrt{(\gamma + 1)/2} (x - st)/l}$$

и предполагаются постоянными перед фронтом скачка ($p = p_2, q = 0$) и за скачком ($p = p_1, q = 0$). На рис. 12.1 показаны графики $V, p + q$ и q для типичного случая.



То обстоятельство, что соотношения Гюгонио точно удовлетворяются данными решениями, означает, что этот «размазанный» скачок распространяется с точно такой же скоростью, как и разрывной скачок для заданного давления p_1 за фронтом, и

Р и с. 12.1. Свойства стационарного скачка с учетом псевдовязкости.

Типичные решения, соответствующие формуле (12.44) и следующим за ней формулам: профили удельного объема, полного давления и псевдовязкого давления.

порождает точно такое же возрастание энтропии. (Это не является неожиданным, потому что соотношения Гюгонио выводятся из законов сохранения массы, количества движения и энергии, а эти законы сохранения вытекают непосредственно из уравнений (12.40) независимо от предположения о форме q . При этом, конечно, существенно, что одно и то же полное давление $p + q$ входит в уравнение энергии и в уравнение движения.) Следовательно, это решение является приемлемой аппроксимацией действительного скачка, если его толщина, равная приблизительно $2l$ или $3l$, достаточно мала¹⁾.

¹⁾ Подлинная толщина, получающаяся на практике, несколько меньше, чем расстояние $(\sqrt{2/(\gamma + 1)} \pi l)$ между штриховыми линиями, на рис. 12.1; точки сопряжения полувольты синусоиды и прямых, соответствующих постоянным значениям, при численном решении, конечно, незаметны. Если эффективную толщину определять как полное изменение V , деленное на максимальный наклон dV/dW , то для аналитического решения (12.44) получается $\sqrt{2/(\gamma + 1)} 2l$; приблизительно то же самое получается при численном решении.

§ 12.11. Конечноразностные уравнения

Конечноразностные системы можно получать различными способами, например модифицируя уравнения (12.10) включением в них членов с псевдовязкостью. В этих членах мы заменим l на $a\Delta x$, где a — безразмерная постоянная, заключенная между 1,5 и 2,0, что дает толщину скачка в 3—5 шагов Δx . Если шаг меняется, то в местах с более крупной сеткой это приведет к большей толщине скачка.

П. Лакс заметил [1954, стр. 165], что приравнивание $l = a\Delta x$ влечет (или по меньшей мере подсказывает) некоторое фундаментальное несоответствие. Предельный переход при Δt и $\Delta x \rightarrow 0$ предположительно приводит к точному решению дифференциальных уравнений (12.40) при фиксированном l , а решение уравнений (12.40) предположительно стремится к точному решению исходных уравнений гидродинамики (включая соотношения Гюгонио) при $l \rightarrow 0$; но нет видимых аргументов для утверждения, что получится тот же самый результат, когда Δt , Δx и l стремятся к нулю совместно при $l = a\Delta x$. Конечно, при реальных вычислениях предельный переход не выполняют до конца, а удовлетворяются некоторым множеством значений a , Δt , Δx ; тем не менее вопрос, поставленный Лаксом, вполне уместен, потому что если при предельном переходе в самом деле необходимо полагать Δt , $\Delta x \rightarrow 0$ сначала при фиксированном l , то можно было бы заключить, что численное решение может быть точным только при $a \gg 1$. Данные, которыми мы располагаем (см. § 12.13), показывают, однако, что это решение может быть весьма точным для значений a , равных 1 или 2, но это в известном смысле следует считать некоторым счастливым совпадением.

Наиболее часто применяются такие разностные уравнения:

$$\begin{aligned} \frac{R_j^{n+1} - R_j^n}{\Delta t} &= u_j^{n+1}, \\ \frac{u_j^{n+1} - u_j^n}{\Delta t} &= -\frac{1}{\rho_0} \frac{(\delta p)_j^n + (\delta q)_j^n}{\Delta r} \left(\frac{R_j^n}{r_j} \right)^{\alpha-1}, \end{aligned} \quad (12.45)$$

$$V_{j+1/2}^{n+1} = \frac{1}{\rho_0} \frac{(R_{j+1}^{n+1})^\alpha - (R_j^{n+1})^\alpha}{(r_{j+1})^\alpha - (r_j)^\alpha},$$

$$\mathcal{E}_{j+1/2}^{n+1} - \mathcal{E}_{j+1/2}^n + \left(\frac{p_{j+1/2}^{n+1} + p_{j+1/2}^n}{2} + q_{j+1/2}^{n+1} \right) (V_{j+1/2}^{n+1} - V_{j+1/2}^n) = 0,$$

$$p_{j+1/2}^{n+1} = f(\mathcal{E}_{j+1/2}^{n+1}, V_{j+1/2}^{n+1}),$$

$$q_{j+1/2}^n = \begin{cases} \frac{2a^2}{V_{j+1/2}^n + V_{j+1/2}^{n-1}} [(\delta u)_{j+1/2}^n]^2, & \text{если } (\delta u)_{j+1/2}^n < 0, \\ 0, & \text{если } (\delta u)_{j+1/2}^n \geq 0. \end{cases}$$

Здесь $\alpha = 1, 2, 3$ для случаев симметрии слоя, цилиндрической и сферической симметрии соответственно.

Нужно отметить, что $q_{j+1/2}^n$ подобно u_j^n в действительности центрировано по времени при $t = t^{n-1/2}$, так что член с q в четвертом из уравнений (12.45) центрирован правильно, а во втором уравнении неправильно. Однако не стоит переписывать второе уравнение, центрируя его: действительно, поступая так, можно прийти к неустойчивости, если не делать это частным и не очень удобным способом. Быть может, можно оправдать неправильное центрирование q тем, что q не является физической величиной, имеющей непосредственное отношение к делу, и поэтому может трактоваться в разностных уравнениях не так точно, как остальные величины.

Иногда выгоднее выражать q через $\partial V / \partial t$, нежели через $\partial u / \partial r$. Соответствующая формула имеет вид

$$q = \frac{(\rho_0 l)^2}{V} \left(\frac{\partial V}{\partial t} \right)^2,$$

что в точности эквивалентно выражению (12.41) в одномерном случае и имеет такой же главный член, как и выражение (12.41), в случае цилиндрической или сферической симметрии. Но теперь вместо $l = a\Delta r$ мы должны писать $l = a\Delta r (r/R)^{\alpha-1}$ в силу следующих соображений: мы не имеем аналитического решения, которым можно было бы руководствоваться в случае цилиндрической или сферической симметрии, но если толщина скачка мала по сравнению с расстоянием от начала отсчета, то, вероятно, будет приближенно пригодной теория плоского скачка.

Действительная толщина скачка равна, конечно, интервалу δR зависимого переменного R , захватываемому скачком, а не соответствующему интервалу δr лагранжевой координаты. Теория плоского скачка дает, что δr пропорционально l и поэтому в плоском случае δR должно быть пропорциональным $l \partial R / \partial r = l \rho_0 / \rho$. Следовательно, δR будет приближенно пропорциональным $l \rho_0 / \rho$ также и в неплоских случаях.

Однако требуется не постоянство длин интервалов δR , захватываемых скачками, а постоянство числа узлов сетки, захватываемых ими (примерно 3—5 узлов), следовательно, мы полагаем $\delta r = a\Delta r$, где a , как и ранее, — безразмерная величина,

т. е. полагаем

$$a \Delta r = \delta r = \frac{\delta R}{\partial R / \partial r} = \frac{l \rho_0 / \rho}{(r/R)^{\alpha-1} (\rho_0/\rho)},$$

что дает результат, установленный для l . Поэтому мы заменим последнее уравнение (12.45) уравнением

$$q_{j+1/2}^n = \begin{cases} \frac{2(a\rho_0 \Delta r)^2}{V_{j+1/2}^n + V_{j+1/2}^{n-1}} \left(\frac{V_{j+1/2}^n - V_{j+1/2}^{n-1}}{\Delta t} \right)^2 \left(\frac{r_{j+1}}{R_{j+1}^n} \right)^{2\alpha-2}, & \text{если } V_{j+1/2}^n - V_{j+1/2}^{n-1} < 0, \\ 0 & \text{в противном случае.} \end{cases} \quad (12.46)$$

Множитель $(r_{j+1}/R_{j+1}^n)^{2\alpha-2}$ будет играть важную роль в многомерных задачах. При желании этот множитель можно центрировать по времени так:

$$\left(\frac{2r_{j+1}}{R_{j+1}^n + R_{j+1}^{n-1}} \right)^{2\alpha-2},$$

но нет оснований для его центрирования по пространству, потому что он уже имеет второй порядок точности, если r и R взяты в одной и той же точке пространства. Мы написали $j+1$ вместо j , чтобы устранить неопределенность в начале отсчета, где $r_0 = R_0^n = 0$.

Отсутствие аналогичного множителя в последнем из уравнений (12.45) можно усмотреть, замечая, что замена в исходной формуле (12.45) $(\delta R)^2$ через $(\delta r)^2$ в точности компенсируется заменой $(\partial u / \partial R)^2$ через $(\partial u / \partial r)^2$.

Если применяются переменные Эйлера, то соответственно полагают

$$q = \begin{cases} l^2 \rho \left(\frac{\partial u}{\partial R} \right)^2 & \text{при } \frac{\partial u}{\partial R} < 0, \\ 0 & \text{в противном случае} \end{cases}$$

и $l = a\Delta R$, так что в разностных уравнениях

$$q_{j+1/2}^n = \begin{cases} 1/2 a^2 (\rho_{j+1/2}^n + \rho_{j+1/2}^{n-1}) (u_{j+1}^n - u_j^n)^2 & \text{при } u_{j+1}^n - u_j^n < 0, \\ 0 & \text{в противном случае.} \end{cases} \quad (12.47)$$

Это уравнение имеет по существу ту же форму, что и последнее уравнение (12.45), но вследствие отличного значения нижних индексов приводит к иному результату: действительная толщина скачка равна теперь не δr , а δR , т. е. постоянна.

Многими авторами были предложены и опробованы другие способы введения искусственной вязкости, но мы не будем на них останавливаться, поскольку нам кажется, что все они уступают методу Лакса — Вендроффа, который будет описан в § 12.14.

§ 12.12. Устойчивость конечноразностных уравнений

Для исследования устойчивости уравнений (12.45) применяется обычный интуитивный подход: выписываются уравнения в вариациях для разностных уравнений, коэффициенты рассматриваются как (почти) постоянные (в любой малой области) и применяется теория, изложенная в гл. 4. В случае симметрии слоя с простым уравнением состояния $\mathcal{E} = pV/(\gamma - 1)$ эта программа может быть доведена до получения пробного условия устойчивости. Однако возникает еще одна трудность в связи с неизбежным существенным изменением коэффициентов (полученных после линеаризации) на расстояниях, сравнимых с Δx ; влияние этого обстоятельства рассматривается в конце настоящего параграфа и в следующем параграфе. Удобно брать u , V , p и q в качестве зависимых переменных, так что матрица перехода $G(\Delta t, k)$ будет иметь четвертый порядок. Предположим, что мы имеем дело с областью плоскости x, t , в которой всюду $\partial u / \partial x \leq 0$, как в примере со стационарным скачком, рассмотренным в § 12.10. (В области, для которой $\partial u / \partial x \geq 0$, член с псевдовязкостью отсутствует, и условие устойчивости такое же, как и в § 12.5.)

Заметим, что мы допускаем, что изменение u не меняет знака $\delta u_{j+1/2}^n$ в последнем уравнении (12.45); это, очевидно, верно для точек, в которых член нулевого порядка $\delta u_{j+1/2}^n$ отрицателен, потому что первая вариация по определению предполагается меньшей, чем любая не обращающаяся в нуль величина нулевого порядка. Если же величина нулевого порядка равна нулю, то вариация q^n в любом случае обращается в нуль с точностью до первого порядка. Можно установить, что два из собственных значений матрицы $G(\Delta t, k)$ равны $\lambda = 1$ и $\lambda = 0$ и не зависят от Δt и k . Другие два собственных значения удовлетворяют квадратному уравнению

$$(\lambda - 1) \left\{ \lambda - 1 - \frac{\lambda + 1}{2} (\gamma - 1) z \Delta t \right\} + \frac{4\Delta t}{(\Delta x)^2} \sin^2 \beta \left[\left(2(\lambda - 1) + \frac{\lambda + 1}{2} z \Delta t \right) l^2 z \left(1 + \frac{\gamma - 1}{2} z \Delta t \right) + \lambda \mu \Delta t \right] = 0,$$

где введены обозначения

$$\mu = [\gamma p + (\gamma - 1) q] / V \rho_0^2, \\ z = -\frac{1}{V \rho_0} \frac{\partial u}{\partial x} = -\frac{1}{V} \frac{\partial V}{\partial t} > 0$$

(это — величины нулевого порядка, играющие роль коэффициентов в уравнениях для вариаций; их следует трактовать как локально постоянные, и поэтому их верхние и нижние индексы опущены) и, как и ранее, $\beta = k\Delta x/2$, а $l = a\Delta x$ — множитель, входящий в коэффициент при q в уравнении (12.41). Из квадратного уравнения для λ очевидно следует, что для устойчивости при $\Delta t, \Delta x \rightarrow 0$ и фиксированном l в любом случае требуется, чтобы $\Delta t = O[(\Delta x)^2]$. Поэтому если мы положим $\Delta t/(\Delta x)^2 = \text{const}$ и устремим Δt и Δx к нулю, то корни квадратного уравнения в пределе будут равны

$$\lambda = 1 \quad \text{и} \quad \lambda = 1 - 8l^2 z [\Delta t/(\Delta x)^2] \sin^2 \beta.$$

Так как z положительно, то для устойчивости требуется выполнение неравенства

$$8l^2 z [\Delta t/(\Delta x)^2] \leq 2. \quad (12.48)$$

В первом издании приводились рассуждения с целью показать, что с точностью первого порядка корни этого квадратного уравнения равны

$$\lambda = 1 + O(\Delta t) \quad \text{и} \quad \lambda = 1 - 8l^2 z [\Delta t/(\Delta x)^2] \sin^2 \beta + O(\Delta t).$$

Тогда неравенство (12.48) гарантировало бы выполнение условий фон Неймана. Теперь эти рассуждения кажутся сомнительными частично из-за ошибки в квадратном уравнении, допущенной в первом издании и обнаруженной Дж. Уайтом из Лос-Аламоса, который детально изучал вопрос об устойчивости, и частично из-за неполного учета всех членов первого порядка. Во всяком случае такое рассмотрение устойчивости дает лишь приближенную ориентировку в этом вопросе (причины этого будут изложены в следующем абзаце) и должно быть дополнено экспериментальными исследованиями, подобными описанным в § 12.13. Поэтому мы ограничимся членами нулевого порядка и примем неравенство (12.48) в качестве пробного условия устойчивости.

Независимо от нестрогости вывода смысл неравенства (12.48) не вполне ясен в силу следующих соображений: понятие устойчивости связано с поведением решения при $\Delta t, \Delta x$, стремящихся к нулю, и начинает приобретать смысл только при Δt и Δx значительно меньших, чем все времена и расстояния, характеризующие данный физический процесс, в частности при Δx , значительно меньших, чем толщина скачка, т. е. при $\Delta x \ll l$. Однако Δx никогда не делается столь малым, потому что мы всегда преднамеренно выбираем l сравнимым с Δx . В силу этих соображений неравенству (12.48) не хватает не только строгости вывода, но и точности; однако оно может служить грубым руководством для выбора шага.

Можно подумать, что в вопросе об устойчивости получится лучшее приближение, если положить $l = a\Delta x$ и считать $\Delta t/\Delta x$ и a фиксированными при $\Delta t, \Delta x \rightarrow 0$. Это неудобно тем, что в области скачка величины нулевого порядка p, q, V и u будут представляться все более и более крутыми графиками, так что даже при самом богатом воображении их нельзя считать локально почти постоянными, но это дает правильный результат в области между скачками. Таким путем получается условие

$$\sqrt{\gamma p V} \Delta t / (\rho_0 V \Delta x) \leq 1, \quad (12.49)$$

в котором мы узнаем условие Куранта — Фридрихса — Леви, потому что $c = \sqrt{\gamma p V}$ — адиабатическая скорость звука, а $\Delta X = \rho_0 V \Delta x$ — расстояние между двумя точками, лагранжевы координаты которых отличаются на Δx . Мы видим, таким образом, что введение членов с псевдовязкостью не изменяет устойчивости обычных разностных уравнений для гладкого течения жидкости.

При использовании условия (12.48) на практике нужно иметь оценку для величины нулевого порядка

$$z = -\frac{1}{V\rho_0} \frac{\partial u}{\partial x} = -\frac{1}{V} \frac{\partial V}{\partial t},$$

и для получения этой оценки мы воспользуемся аналитическим решением (12.44) для плоского стационарного скачка. Полученное таким образом значение z изменяется при переходе через область скачка, и, очевидно, мы должны использовать максимум z , который находится без особого труда. Полагая $l = a\Delta x$, мы можем придать условию (12.48) следующий вид:

$$\frac{s \Delta t}{\Delta x} \leq \frac{1}{2a} \sqrt{\frac{2\eta}{\gamma+1}} \frac{1}{\eta-1}, \quad (12.50)$$

где s — скорость скачка (точнее, производная по времени от лагранжевой координаты скачка), а $\eta = V_2/V_1$ — коэффициент сжатия объема при переходе через скачок.

Чтобы сравнить два условия (12.49) и (12.50), которые применяются вне и внутри скачка соответственно, обозначим через p_f и V_f давление и удельный объем непосредственно за скачком¹⁾ и положим

$$L_f = \sqrt{\gamma p_f V_f} \frac{\Delta t}{\rho_0 V_f \Delta x};$$

L_f часто называют «числом Куранта». После несложных преобразований неравенства (12.50) с помощью соотношений Гюго-

¹⁾ Очевидно, что если условие (12.49) выполняется за скачком, то оно будет выполняться также и перед скачком, так как скорость звука перед скачком меньше, а удельный объем больше, чем за скачком.

нию можно записать эти два условия следующим образом:

$$L_f \leq 1, \\ L_f \leq \frac{V \eta [\eta - (\gamma - 1)(\gamma + 1)]}{2a(\eta - 1)}. \quad (12.51)$$

Для достаточно слабых скачков ($\eta - 1$ мало) Δt ограничивается первым из этих неравенств, а для сильных скачков более существенным является второе. Если нельзя оценить интенсивность скачка, то можно просто предполагать худшее: наиболее ограничительное условие для L_f получается при бесконечно сильном скачке, для которого $\eta = (\gamma + 1)/(\gamma - 1)$, и тогда мы находим, что

$$L_f \leq V\sqrt{\gamma}/2a. \quad (12.52)$$

Соответственно этому для задач с очень сильными скачками метод псевдовязкости уменьшает допустимые Δt примерно на множитель $V\sqrt{\gamma}/(2a)$, который практически приблизительно равен $1/3$. Для прочих задач это уменьшение допустимых Δt оказывается меньше и определяется неравенствами (12.51).

[Дж. Уайт из Лос-Аламоса сообщил нам о следующем часто используемом условии устойчивости для метода псевдовязкости:

$$L + 4a^2 |V_{j+1/2}^{n+1} - V_{j+1/2}^n| / V_{j+1/2}^n < 1,$$

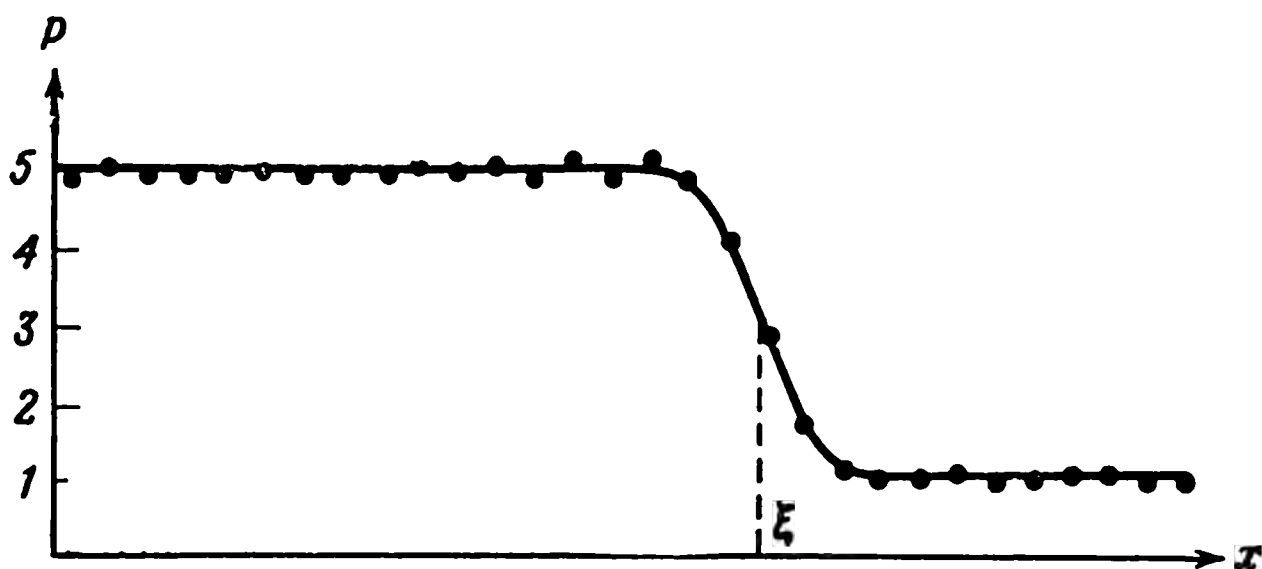
где через L обозначена левая часть неравенства (12.49), а второй член является разностной аппроксимацией половины левой части неравенства (12.48).]

Как уже было отмечено выше, этим результатам не хватает не только строгости вывода, но и точности содержания. Численные эксперименты, описанные в следующем параграфе, наводят на мысль, что действительное ограничение L_j для сильных скачков несколько слабее второго неравенства (12.51), а для слабых скачков несколько сильнее его.

§ 12.13. Численная проверка метода псевдовязкости

Численная проверка метода осуществлялась различными исследователями разными способами. В одной серии расчетов, проведенной на машине УНИВАК, использовались уравнения (12.45) при $\alpha = 1$ и при $f(\mathcal{E}, V) = (\gamma - 1)\mathcal{E}/V$. Можно считать, что в этом случае уравнения описывают движение идеального газа в трубе. Газ сначала покоится под давлением p_0 , а затем на одном конце трубы прикладывается давление Πp_0 , где $\Pi > 1$, как в случае поршня под постоянным давлением.

Машина рассчитывала движение газа в отрезке трубы, начинающемся с поршня и представленном сеткой из 30 точек. Давление как функция x для этих 30 точек в некоторый момент времени может быть таким, как показано на рис. 12.2, где наличие скачка очевидно. С течением времени этот скачок движется вправо, и машина автоматически определяла его положение $x = \xi(t)$, находя точку, в которой давление равно $(p_0 + \Pi p_0)/2$; эта точка находилась как обращение линейной интерполяции значений p в узлах сетки. Когда скачок достигал



Р и с. 12.2. Профиль давления для стационарного скачка, полученный из разностных уравнений; $a^2 = 4,0$; $\Pi = 5,0$; $\gamma = 2,0$; $L_f = 0,2236$.

20-го узла сетки, машина автоматически прибавляла 10 узлов сетки с внешним давлением p_0 впереди скачка и снимала 10 узлов сетки, примыкающих к поршню. Таким образом, за движением скачка можно следить неограниченно долго, хотя в каждый момент времени при вычислениях рассматривается только 30 узлов сетки.

К началу работы машины в нее должны быть введены значения Π , γ , a^2 и Δt . Затем машина автоматически следила за скачком, как это было описано выше. После некоторого начального периода, достаточного для того, чтобы скачок принял стационарную форму и ушел по меньшей мере на 10 узлов сетки от поршня, машина начинала определять его положение $\xi(t)$ и другие величины для следующих 100 шагов, т. е. для времени $100 \Delta t$.

Машина вычисляла скорость скачка путем подгонки $\xi(t)$ как функции от t к прямой по методу наименьших квадратов, удельный объем V и скорость u частиц жидкости за скачком путем усреднения по последним пяти (или десяти) точкам, а также толщину скачка для оценки максимума наклона давления как функции x . Затем она печатала краткую таблицу, содержащую значения этих величин (определенные, как описано выше), их вероятные погрешности (определенные по отклонениям от точных значений на протяжении 100 шагов), а также

их теоретические или точные значения (определенные из уравнений Ренкина — Гюгонио). Машина также записывала более подробные таблицы на магнитную ленту для последующего вывода на печатающее устройство, управляемое магнитной лентой (эти таблицы содержали полные профили давления, удельного объема, скорости частиц жидкости, энергии и энтропии для последних пяти шагов), а затем ожидала ввода параметров следующей задачи.

Рис. 12.2 показывает профиль давления для типичного шага задачи, в которой отношение давлений Π было равно 5,0, а γ и a равнялись 2,0. Таблица, напечатанная машиной при решении этой задачи, воспроизведена в виде табл. 12.1. (Напечатанная машиной таблица здесь несколько преобразована для того, чтобы привести обозначения в соответствие с принятыми в тексте.)

Таблица 12.1

Машинный расчет скачка

Дата печатания	17 ноября 1953 г.		
0,01 γ	002000000000		
0,01 Π	005000000000		
0,01 a^2	004000000000		
L_0	005000000000		
Величина	Теорет. значение	Эксперим. значение	Вероятн. погрешность
Скорость скачка	0,1414213	0,1414191	0,0000147
V (5 точек)	0,5000000	0,5025913	0,0000925
V (10 точек)		0,5013867	0,0000490
u (5 точек)	1,4142135	1,4143566	0,0001610
u (10 точек)		1,4141951	0,0001304
интервал сетки		3,3651549	0,0023169
L_f	0,2236067		
L_f max	0,3535533		
Следующий шаг			
0,01 γ и т. д.			

Скорость скачка измерялась в интервалах сетки за шаг, т. е. печаталась величина

$$\frac{\Delta t}{\Delta x} \frac{d\xi(t)}{dt}.$$

Ширина скачка выражалась через интервалы сетки, а удельный объем V и скорость частиц жидкости u за скачком — в некоторых единицах, связанных с удельным объемом и скоростью

звука перед скачком. Как и в предыдущем параграфе, L_f означает величину числа Куранта за скачком, а $L_{j\max}$ — величину, которую L_f должно иметь для того, чтобы вычисления велись на границе неустойчивости в соответствии с критерием (12.51). Мы видим, что в этой задаче теоретические и экспериментальные значения согласуются очень хорошо (согласование значе-

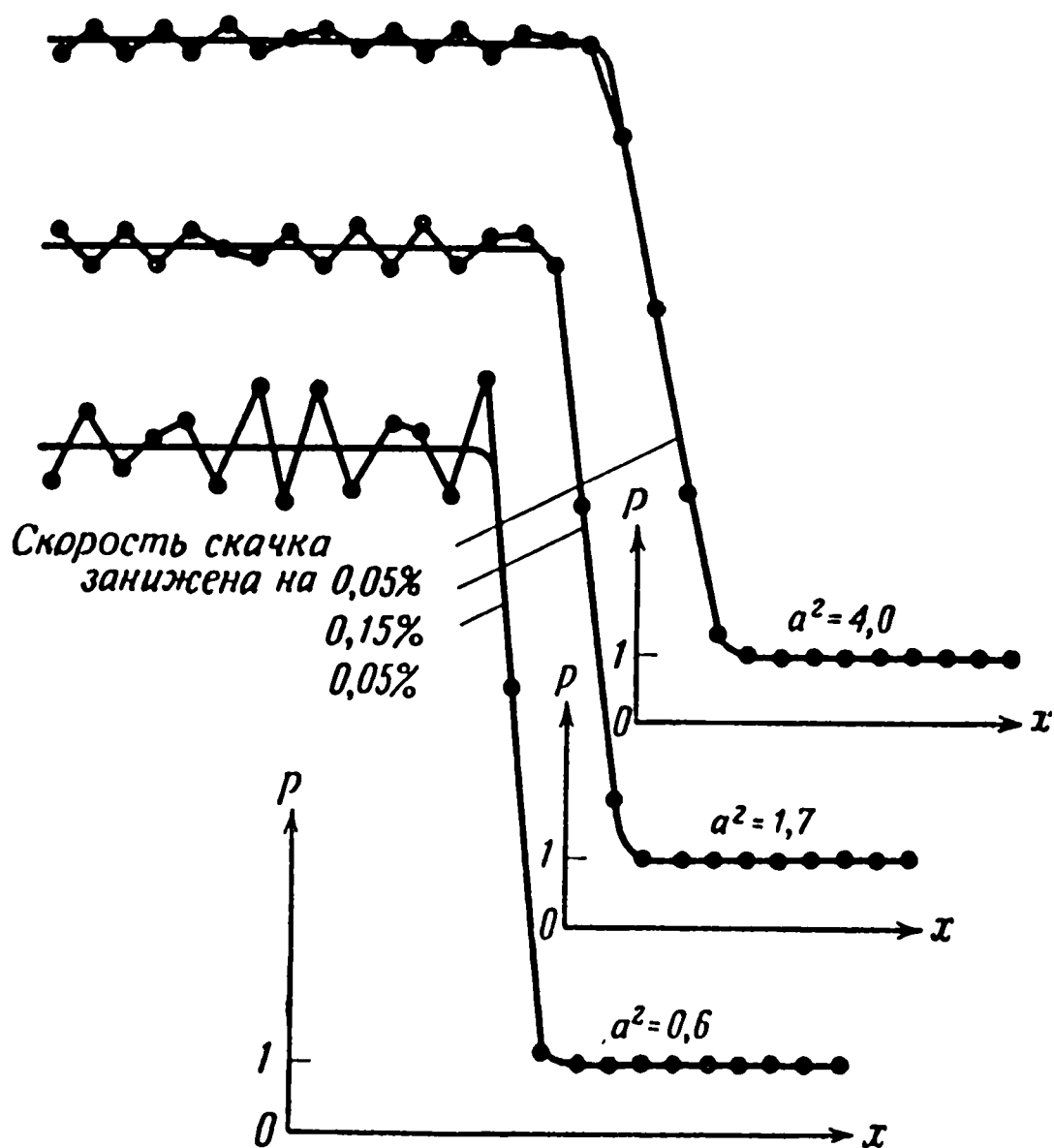


Рис. 12.3. То же, что на рис. 12.2, но для более сильного скачка и для трех различных значений коэффициента псевдовязкости; $\gamma = 2,0$; $\Pi = 10,0$; $L_f = 0,3453$.

ний скорости скачка в данном частном случае случайно оказывается очень хорошим).

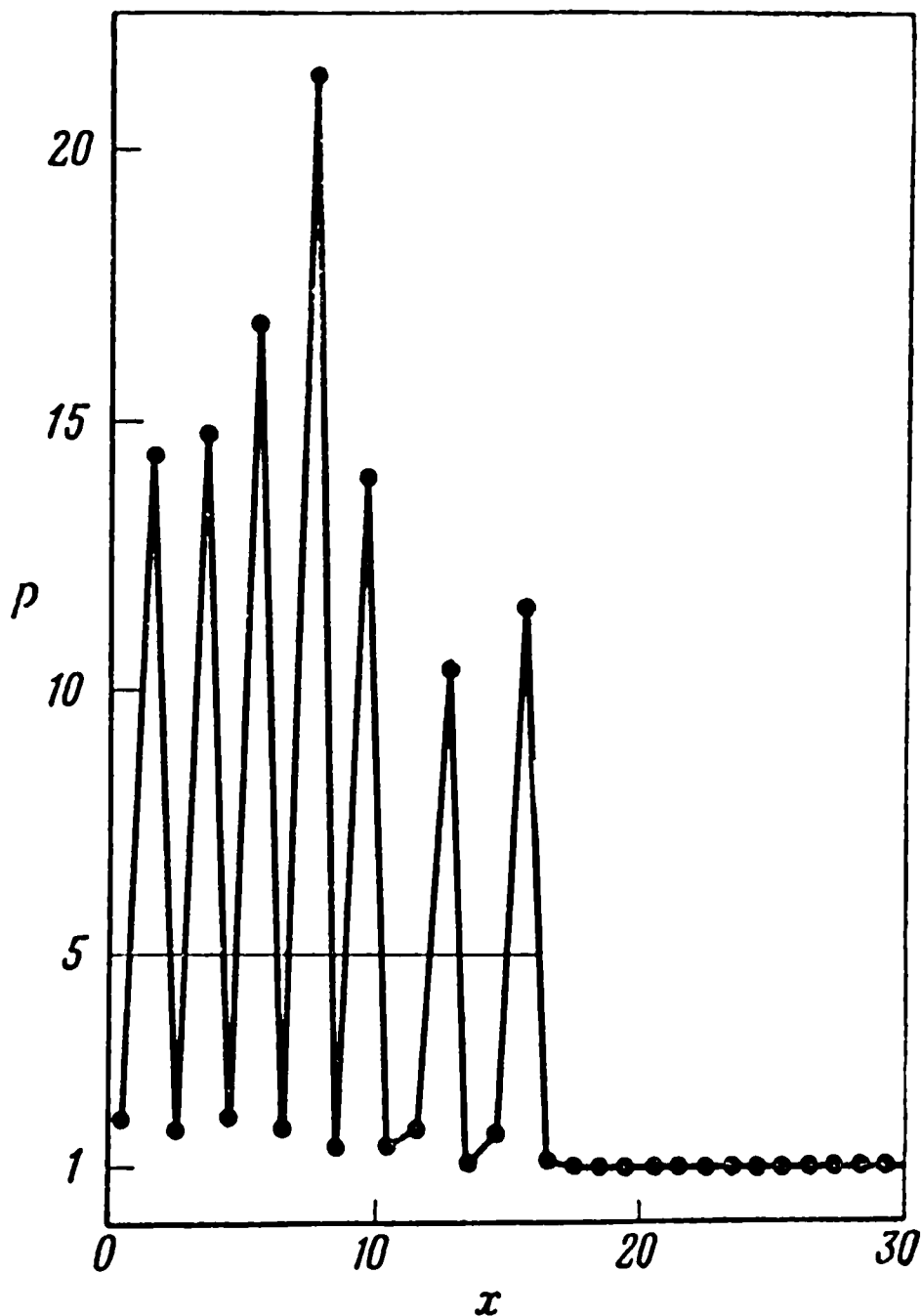
По этой программе было проведено несколько десятков расчетов при различных значениях параметров. Во всех случаях, когда постоянная a , характеризующая псевдовязкость, была не меньше чем примерно $3/4$, данные «теории» и «эксперимента» согласовывались исключительно хорошо. На рис. 12.3 изображены профили давления для типичного шага в каждой из трех задач, отличающихся значением a . Для этих задач Π было взято равным $10,0$, γ было взято равным $2,0$, а Δt было выбрано так, что число Куранта за скачком равнялось примерно $1/3$. При большом значении коэффициента псевдовязкости получалась большая толщина скачка и гладкий профиль давления за скачком, в то время как при малом значении коэффициента

псевдовязкости скачок оказывался более крутым, но за ним возникали колебания давления. При выборе значения a нужно идти на компромисс между желаемой крутизной профиля давления на скачке и гладкостью профиля за ним. В этих трех задачах скорость скачка с точностью до 0,1% совпадает со значением, найденным из уравнений Гюгонио.

Рис. 12.4 показывает, что происходит, когда член с псевдовязкостью опускается (коэффициент полагается равным нулю). За скачком получаются резкие колебания давления, и скорость скачка определяется с ошибкой в 10%.

Быть может, нужно отметить, что эти колебания не означают неустойчивости. Вычисления, результаты которых приве-

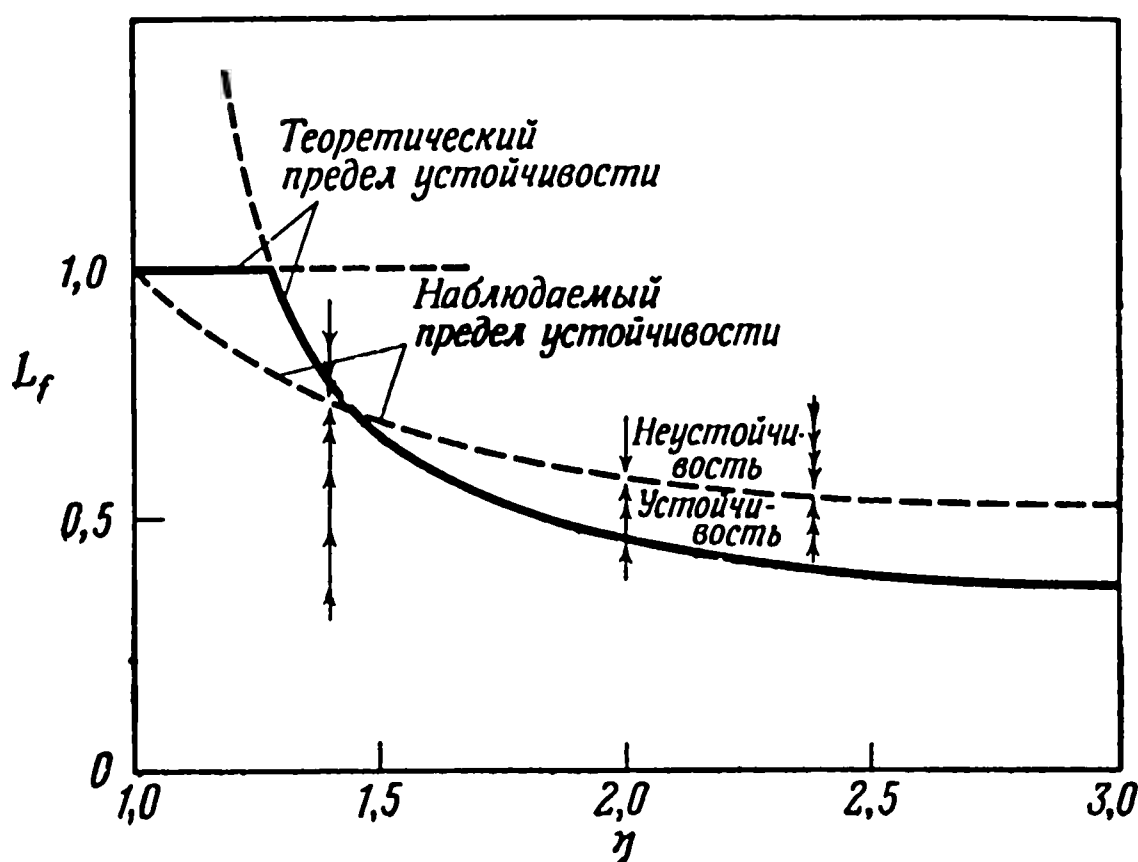
Рис. 12.4. То же, что на рис. 12.2, но при полном отсутствии псевдовязкости; $a^2 = 0$; $\Pi = 5,0$; $\gamma = 2,0$; $L_f = 0,2236$; скорость скачка завышена на 10,4%.



дены на рис. 12.4, могут быть продолжены сколь угодно долго, причем амплитуда колебаний будет оставаться примерно такой же, как показано. Этим колебаниям в действительности может быть дано физическое истолкование. Конечноразностные уравнения приближенно описывают движение некоторой модели жидкости. В этой модели масса сконцентрирована в нескольких тонких параллельных пластинках, пространство между которыми заполнено невесомой сжимаемой жидкостью, служащей упругим элементом. При движении скачка пластинки приходят в движение, подобно молекулам газа (одномерного газа в данной задаче). Р. Пайерлс указал (неопубликованная работа, Лос-Аламос), что эти колебания аналогичны тепловому возбуждению молекул и что они представляют внутреннюю энергию, которая должна появиться в жидкости после прохождения скачка согласно уравнениям Гюгонио. Целью введения

псевдовязкости является превращение этой энергии колебаний в действительно внутреннюю энергию.

Программа для УНИВАКа, описанная выше, применялась также для изучения устойчивости разностных уравнений. Задача решалась при заданных значениях Π , γ и a и при различных значениях Δt . Если при выбранном значении Δt решение задачи шло гладко и нормально для 100 шагов и давало согласование между «теорией» и «экспериментом» с точностью



Р и с. 12.5. Результаты численного исследования устойчивости разностных уравнений с учетом псевдовязкости; $\gamma = 2,0$; $a^2 = 4$; $\eta = V_i/V_f$; $L_f = \eta c \Delta t/\Delta x$; c — адиабатическая скорость звука за скачком.

до 1% или лучше, то уравнения с этим значением Δt считались устойчивыми. Во всех остальных случаях машина останавливалась из-за переполнения, потому что некоторая величина выходила за предела шкалы задолго до завершения 100 шагов. В этих случаях уравнения считались неустойчивыми.

Рис. 12.5 показывает результат исследований такого рода при $\gamma = 2,0$ и $a = 2,0$. Изучались три скачка различной интенсивности, соответствующей значениям коэффициента сжатия $\eta = 1,4, 2,0$ и $2,383^1)$. В каждом из этих случаев конец стрелки, направленной вверх, показывает значение L_f , для которого уравнения устойчивы, а конец стрелки, направленной вниз, показывает значение L_f , для которого они неустойчивы. Пунктирная линия была проведена посередине между стрелками, чтобы дать приближенно наблюдаемый предел устойчивости. Сплошная линия изображает теоретический предел устойчивости, оп-

¹⁾ При $\gamma = 2,0$ наибольшее возможное значение коэффициента сжатия η равно 3,0.

ределяемый неравенствами (12.51). Мы видим, что уравнения несколько более устойчивы для сильных скачков и несколько менее устойчивы для слабых скачков, чем можно было ожидать.

§ 12.14. Метод Лакса — Вендроффа для расчета движения скачков

Введение искусственной вязкости приводит к затуханию высокочастотных компонент решения в процессе вычислений. Любую разностную схему, для которой такие компоненты затухают, можно считать схемой с включением вязкости, даже если в ней нет специальной группы членов, которые можно было бы считать представляющими вязкость. Обычно уравнения такой схемы можно переписать так, что диссипативные члены появляются как поправка к другой схеме, не подавляющей высокочастотные компоненты.

Схему Лакса — Вендроффа (12.22) для уравнения

$$\frac{\partial U}{\partial t} + A \frac{\partial U}{\partial x} = 0, \quad (12.53)$$

где U — вектор, а A — постоянная матрица, можно представить в виде

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} + A \frac{U_{j+1}^n - U_{j-1}^n}{2 \Delta x} = \frac{\Delta t}{2} A^2 \frac{U_{j+1}^n - 2U_j^n + U_{j-1}^n}{(\Delta x)^2}, \quad (12.54)$$

и может показаться заманчивым рассматривать правую часть этого уравнения как аппроксимацию диссипативного члена $1/2 \Delta t A^2 \partial^2 U / \partial x^2$, который поэтому мог бы быть включен и в дифференциальное уравнение. Но это было бы некорректно, потому что так приписанный диссипативный член имеет тот же порядок, что и погрешность аппроксимации левой части; в таком случае можно было бы переносить члены из одной части в другую и получать множество поправок к дифференциальному уравнению, в том числе и недиссипативных. Для того чтобы сделать выбор диссипативного члена однозначным, предположим, что $U = U(x, t)$ — точное решение уравнения

$$\frac{\partial U}{\partial t} + A \frac{\partial U}{\partial x} = QU, \quad (12.55)$$

и постараемся определить оператор Q так, чтобы

$$U_j^{n+1} - U_j^n + \frac{1}{2} \frac{A \Delta t}{\Delta x} (U_{j+1}^n - U_{j-1}^n) - \frac{1}{2} \left(\frac{A \Delta t}{\Delta x} \right)^2 (U_{j+1}^n - 2U_j^n + U_{j-1}^n) = O[(\Delta t)^4], \quad (12.56)$$

а не $O[(\Delta t)^3]$, как это имеет место при отсутствии члена QU в уравнении (12.55). Считая Q таким дифференциальным оператором с постоянными коэффициентами, что $QU = O[(\Delta t)^2]$, разложим все члены в (12.56) в ряды Тейлора в окрестности точки $j\Delta x$, $n\Delta t$ и приравняем нулю сумму членов порядка $(\Delta t)^3$; используя уравнение (12.55), находим, что

$$QU = -\frac{1}{6} A [(\Delta x)^2 - A^2 (\Delta t)^2] \frac{\partial^3 U}{\partial x^3}. \quad (12.57)$$

К сожалению, дифференциальное уравнение (12.55) с таким оператором QU не диссипативно, хотя оно является дисперсионным, т. е. различные фурье-компоненты точного решения этого уравнения распространяются с различными скоростями, но без изменения амплитуды.

Чтобы получить диссипативный член, необходимо в соотношении (12.56) заменить $O[(\Delta t)^4]$ на $O[(\Delta t)^5]$; тогда Q заменится на $Q + Q'$, где добавочная поправка $Q'U$ дается равенством

$$Q'U = -\frac{1}{8} A^2 \Delta t [(\Delta x)^2 - (A \Delta t)^2] \frac{\partial^4 U}{\partial x^4} \quad (12.58)$$

и вводит диссипацию в (12.55). Это приводит к тому, что амплитуда фурье-компоненты $\exp(ikx)$ уменьшается как $\exp(-\text{const} \cdot k^4 t)$ при условии, что разность $(\Delta x)^2 - (\lambda \Delta t)^2$ положительна для каждого собственного значения λ матрицы A , как это необходимо для устойчивости.

Аналогичный анализ разностных уравнений фон Неймана — Рихтмайера (12.45) показывает, что в этом случае поправка наименьшего порядка (а именно порядка $(\Delta t)^2$) одновременно является как дисперсионной, так и диссипативной.

Иногда приводятся возражения, что уравнения Лакса — Вендроффа должны быть в некотором смысле менее точными, чем центрированные или «петлеобразные» уравнения вследствие затухания таких фурье-компонент $\exp(ikx)$, для которых $k\Delta x \sim 1$, что не имеет места для «петлеобразных» схем и для самих дифференциальных уравнений. Эти возражения отпадут, если учесть фазы фурье-коэффициентов, которые при $k\Delta x \sim 1$ в действительности искажаются как схемой Лакса — Вендроффа, так и петлеобразной схемой на величину порядка $(k\Delta x)^3$, одинаковую для обеих схем и на порядок большую, чем величина искажения (затухания) амплитуд. Нам кажется, что при этих обстоятельствах сохранение неизменными амплитуд высокочастотных фурье-компонент является излишним.

Возможно, следует отметить, что диссипация не приводит к какой-либо потере энергии или других сохраняющихся величин: часть кинетической энергии, соответствующая высокоча-

стотным колебаниям, переходит во внутреннюю энергию жидкости.

Представляется более удобным изучать влияние диссипации на уровне разностных, а не дифференциальных уравнений. Лаксом и Вендроффом [1960] был достигнут известный успех в теоретическом понимании аппроксимаций решений со скачками, полученных с помощью диссипативных разностных уравнений типа Лакса — Вендроффа (их результаты подытожены в конце этого параграфа), но основные доводы в пользу этого метода до сих пор остаются эмпирическими.

В качестве примера мы рассчитывали с помощью уравнений Лакса — Вендроффа распространение плоского стационарного скачка в идеальном газе; серия вычислений аналогична описанной в § 12.13 в связи с методом фон Неймана — Рихтмайера. Была использована двухшаговая схема (12.24а), (12.24б), в которой векторы U и $F(U)$ задавались соотношением (12.21), и уравнение состояния было принято в виде $f(\mathcal{E}, V) = p = (\gamma - 1)\mathcal{E}/V$.

Изучавшаяся краевая задача состояла в том, что поршень вдвигался в трубу с первоначально покоящимся газом постоянной силой, создающей давление, в пять раз большее первоначального. Профиль скорости вблизи фронта скачка для двух случаев показан на рис. 12.6 и 12.7 — оба в тот момент, когда скачок ушел по трубе от поршня на 80 точек сетки (всего было 100 точек). Положение скачка определялось абсциссой, в которой скорость равнялась половине ее конечного значения, а это положение находилось линейной интерполяцией. Скорость скачка определялась по времени его прохождения от 20-й до 80-й точки сетки. Было найдено, что так определенная скорость скачка с точностью до малых долей процента согласуется с теоретической, вычисленной по соотношениям Гюгонио (12.37).

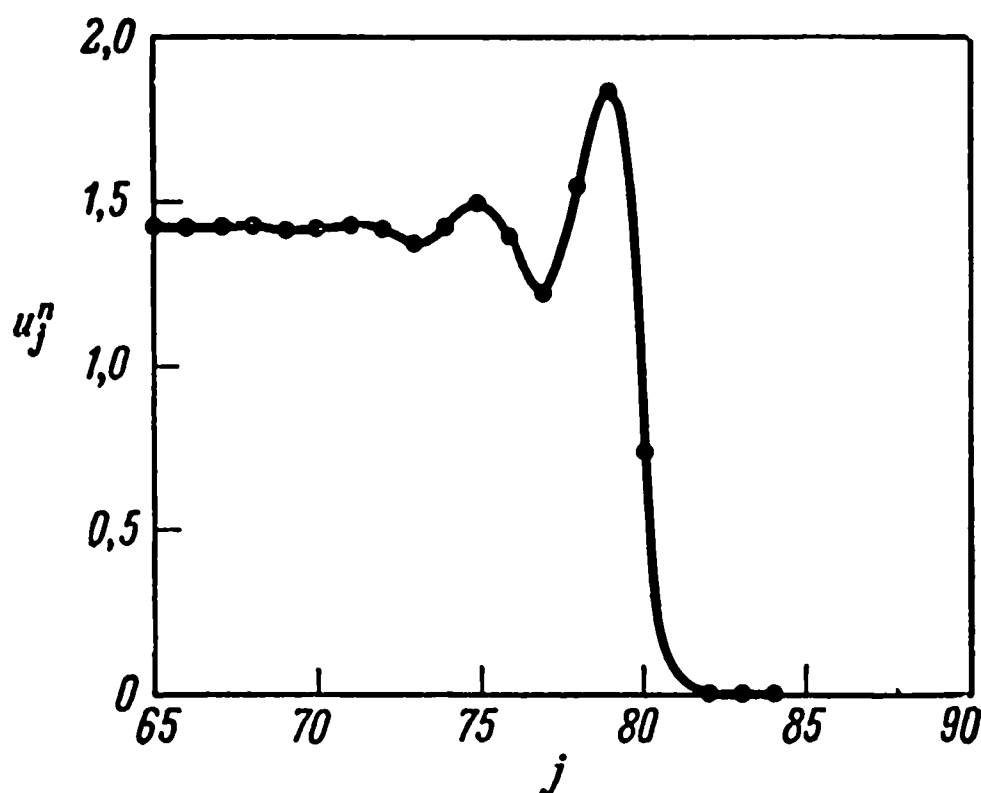
Фронт скачка получается достаточно крутым, но за ним возникают колебания, которые весьма малы, когда число Куранта

$$C_f = \frac{V_0}{V} \frac{c \Delta t}{\Delta x}$$

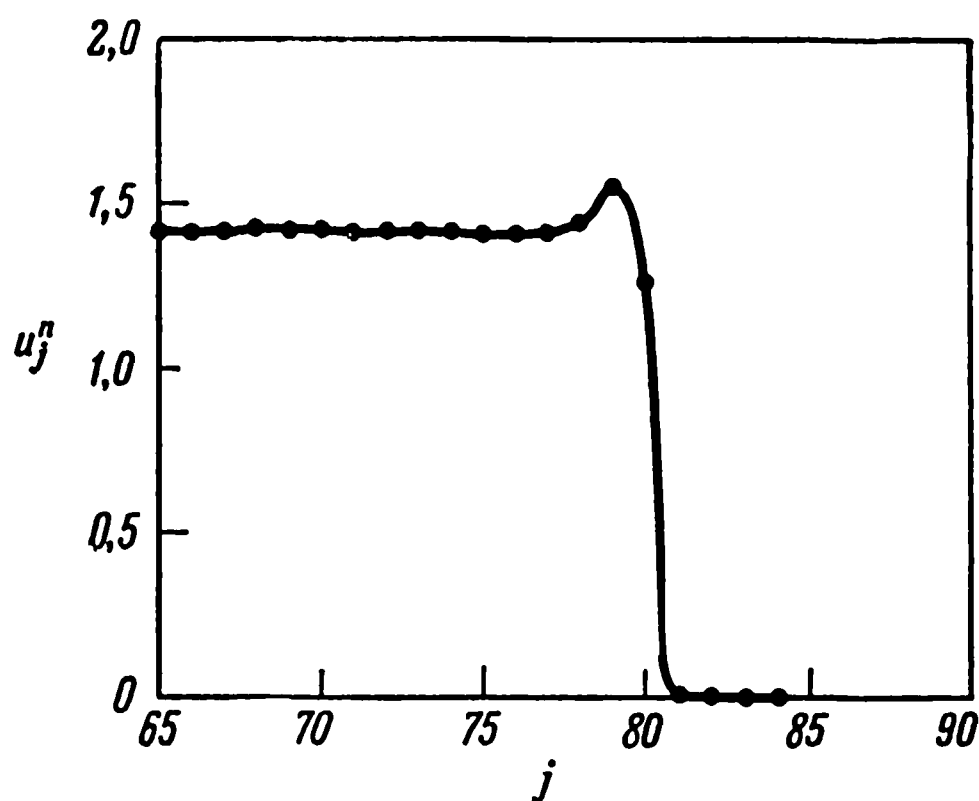
принимает далеко позади скачка максимально допустимое значение $C_f = 1,0$. Эти колебания заметно больше, когда $C_f = 0,5$.

Лакс и Вендрофф [1960] показали, как уменьшить эти колебания путем введения в разностные уравнения диссипативных или вязких членов. Позднее было установлено (Лакс, Бурстейн, Лapidус, частное сообщение), что эти дополнительные члены оказываются необходимыми для предотвращения нелинейной неустойчивости в некоторых задачах. Причина этого состоит в том, что если одно из собственных значений матрицы A равно нулю, то соответствующее собственное значение g матрицы

перехода G , определяемой равенством (12.25), при всех α (т. е. при всех k) равно единице, а не меньше ее для $k > 0$. Иначе говоря, в этом случае система не будет диссипативной в смысле



Р и с. 12.6. Расчет движения скачка по схеме Лакса — Вендроффа; $C_f = 0,5$; $t = 0,28398$;
 $S_{\text{эксп}} = 2,828294$;
 $S_{\text{теор}} = 2,828428$; $P_{\text{кон}}/P_{\text{нач}} = 5,0$; $\gamma = 3,0$.



Р и с. 12.7. То же, что на рис. 12.6, но для большего значения Δt ; $C_f = 1,0$; $t = 0,28398$;
 $S_{\text{эксп}} = 2,828101$;
 $S_{\text{теор}} = 2,828428$; $P_{\text{кон}}/P_{\text{нач}} = 5,0$; $\gamma = 3,0$.

Крайса (см. § 5.4), так что нелинейность может привести к неустойчивости. Собственные значения матрицы A равны u и $u \pm c$, и поэтому возникает опасность появления неустойчивости в точке покоя или в точке звуковой скорости. Предложение Лакса и Вендроффа [1960] состояло в том, чтобы добавить к правой части уравнения (12.22) член, представляющий искусственную вязкость и имеющий вид

$$\frac{\Delta t}{2 \Delta x} [Q_{j+1/2} (U_{j+1} - U_j) - Q_{j-1/2} (U_j - U_{j-1})],$$

где $Q_{j+1/2} = Q_{j+1/2}(U_j, U_{j+1})$ — матрица, которая выбирается так, чтобы получить желаемый эффект диссипации, когда U_j и U_{j+1} значительно отличаются одно от другого, и сделать его

пренебрежимо малым, когда они почти равны. Выбор, сделанный Лаксом и Вендроффом, основывается на следующих соображениях: 1) если матрица A постоянна, то выбором таких новых зависимых переменных в разностных уравнениях (12.54), которые диагонализуют матрицу A , эти уравнения разделяются на p ($p = 3$ в случае уравнений гидродинамики) независимых скалярных уравнений, и естественно добиваться, чтобы диссипация была эффективной для каждого из этих уравнений; 2) для одного скалярного уравнения желательный эффект диссипации и удовлетворительный профиль скачка получаются при

$$Q_{j+1/2}(U_j, U_{j+1}) = \kappa |A(U_{j+1}) - A(U_j)|,$$

где κ — безразмерная постоянная порядка 1, а $A = dF/dU$; теперь F и U — скалярные величины. Чтобы удовлетворить этим требованиям, матрица $Q_{j+1/2}$ должна быть перестановочной с матрицей A , которую в случае переменных коэффициентов мы интерпретируем как $A_{j+1/2} = A(1/2 U_j + 1/2 U_{j+1})$, и i -е собственное значение матрицы $Q_{j+1/2}$ должно равняться $\kappa |a_{j+1}^i - a_j^i|$, где a_j^i и a_{j+1}^i суть i -е собственные значения матриц $A(U_j)$ и $A(U_{j+1})$. Эти соображения приводят к выбору

$$Q_{j+1/2} = g(A_{j+1/2}),$$

где $g(\cdot)$ — многочлен степени $p - 1$ (т. е. степени 2 для уравнений гидродинамики), который должен равняться $\kappa |a_{j+1}^i - a_j^i|$, когда его аргумент равен $a_{j+1/2}^i$, т. е. i -му собственному значению матрицы $A_{j+1/2}$; следовательно, это интерполяционный полином Лагранжа. Коэффициенты этого полинома являются функциями от компонент векторов U_j и U_{j+1} .

Если матрица A постоянна, то $g(\cdot) \equiv 0$, так как величина $|a_{j+1}^i - a_j^i|$ обращается в нуль при любом i ; в общем случае главный диссипативный член (в который входит вторая разность от U) получается заменой Q на константу, что дает $(\Delta t/2\Delta x) Q \delta^2 U$. Поэтому к матрице перехода (12.25) для уравнений Лакса — Вендроффа добавляется дополнительный член $-(\Delta t/\Delta x) Q (1 - \cos \alpha)$; так как собственные значения матрицы Q неотрицательны и, вообще говоря, положительны в области быстрого изменения коэффициентов (такой, как скачок), где нелинейные эффекты наиболее ярко выражены, действительные части (а следовательно, и модули) собственных значений g матрицы G в общем случае уменьшаются даже в тех точках, где собственное значение матрицы A обращается в нуль. Если a и q являются соответствующими собственными значениями матриц A и Q , то собственное значение матрицы G равно

$$g = 1 - i \frac{a \Delta t}{\Delta x} \sin \alpha - \left[\left(\frac{a \Delta t}{\Delta x} \right)^2 + \left(\frac{q \Delta t}{\Delta x} \right)^2 \right] (1 - \cos \alpha).$$

В гладкой части течения $q \ll a$, но в области скачка величина q , равная $\kappa |a_{j+1} - a_j|$, может быть сравнимой с a . Предположим, однако, что даже на скачке q не превосходит κa^0 , где a^0 — максимальное по модулю собственное значение матрицы A . Эллипс в комплексной плоскости, характеризующий g как функцию от α , лежит внутри единичной окружности, если в приведенном выше равенстве выражение в квадратных скобках не превосходит единицы. Поэтому условием устойчивости является неравенство

$$\left(\frac{a^0 \Delta t}{\Delta x}\right)^2 + \kappa \frac{a^0 \Delta t}{\Delta x} < 1,$$

которое для уравнения гидродинамики в эйлеровой форме принимает вид

$$\frac{a^0 \Delta t}{\Delta x} = (|v| + c) \frac{\Delta t}{\Delta x} < \left(1 + \frac{\kappa^2}{4}\right)^{1/2} - \frac{\kappa}{2}.$$

Это условие оказывается более ограничительным, чем условие Куранта, Фридрихса и Леви — например на множитель 0,78 при $\kappa = 0,5$.

Лapidус (частное сообщение) показал, что несколько модифицированная вязкость Лакса — Вендроффа, которая имеет тот же самый главный член $(\Delta t/2\Delta x) Q \delta^2 U$, может быть введена в двухшаговый метод § 12.7 для системы третьего порядка, подобной гидродинамической, где $g(A) = g_0 + g_1 A + g_2 A^2$. Уравнения (12.24а, б) заменяются уравнениями

$$U_{j+1/2}^{n+1/2} = \frac{1}{2}(U_{j+1}^n + U_j^n) - \frac{1}{2}\left(\frac{\Delta t}{\Delta x} + g_2^+\right)(F_{j+1}^n - F_j^n) - \\ - \frac{1}{2} g_1^+(U_{j+1}^n - U_j^n),$$

$$U_j^{n+1} = U_j^n - \frac{\Delta t}{\Delta x}(F_{j+1/2}^{n+1/2} - F_{j-1/2}^{n+1/2}) - \\ - \frac{1}{2} \frac{\Delta t}{\Delta x} [g_0^+(U_{j+1}^n - U_j^n) - g_0^-(U_j^n - U_{j-1}^n)],$$

где g_0^+ , g_1^+ , g_2^+ — коэффициенты полинома, который выражает $Q_{j+1/2}$ через $A_{j+1/2}$, а g_0^- , g_1^- , g_2^- — коэффициенты полинома, выражающего $Q_{j-1/2}$ через $A_{j-1/2}$.

Если применяются зависимые переменные (12.21) в форме Лагранжа, то формулы становятся несколько проще. В этом случае собственные значения матрицы A равны 0 и $\pm c'$, где $c' = V_0 c/V$, а c — адиабатическая скорость звука. Так как c зависит от V и E , то $c' = c'(U)$. Тогда $g(\cdot)$ сводится только

к квадратичному члену и

$$Q_{j+1/2} = \kappa \frac{|c'(U_{j+1}^n) - c'(U_j^n)|}{[c'(U_{j+1/2}^n)]^2} (A_{j+1/2})^2.$$

В качестве некоторого теоретического обоснования того, что эти методы дают правильное описание течений со скачками, мы приведем два результата Лакса и Вендроффа [1960]. Первый из них опирается на понятие слабого решения системы, записанной в дивергентной форме, которое мы сейчас введем. Прежде всего, если $U(x, t)$ — точное решение уравнения (12.16) с заданным начальным условием $U(x, 0) = U^0(x)$, а $W(x, t)$ — любая гладкая векторная функция с тем же числом компонент, что и U (именно p компонент), которая обращается в нуль вне конечной области плоскости x, t (такая функция называется *пробной*), то, умножая скалярно обе части уравнения (12.16) на $W(x, t)$, интегрируя по всем x и по всем $t > 0$ и вычисляя интеграл по частям, находим, что

$$\int \int \left(\frac{\partial W}{\partial t} \cdot U + \frac{\partial W}{\partial x} \cdot F \right) dx dt + \int W(x, 0) \cdot U^0(x) dx = 0.$$

Любая функция $U(x, t)$ независимо от того, является ли она точным решением, удовлетворяющая этому уравнению при всех пробных функциях $W(x, t)$, называется *слабым решением* уравнения (12.16) с начальным значением $U^0(x)$. Из этого определения следует (см. Лакс [1957]), что на разрыве первого рода слабое решение удовлетворяет условию

$$S[U] = [F(U)],$$

где $[f]$ означает разность $f(x+0) - f(x-0)$ (x — координата разрыва), а S — скорость движения разрыва. Так как компоненты U и $F(U)$ — плотность и плотность потоков массы, количества движения и энергии, это условие эквивалентно условиям Ренкина — Гюгонио, если разрыв представляет собой скачок уплотнения, и условиям непрерывности давления и скорости жидкости, если разрыв является контактным. (Это — единственно возможные типы простых разрывов.) Применительно к изложенному выше результаты Лакса и Вендроффа заключаются в следующем.

1. Если решение $U_j^n = U(j\Delta x, n\Delta t)$ уравнений Лакса — Вендроффа (12.22) с вязким членом или без него почти всюду сходится к конечной функции $V(x, t)$, то $V(x, t)$ есть слабое решение дивергентной системы (12.16).

2. В определенном приближении любое стационарное решение разностных уравнений содержит просто переход на протяжении двух интервалов сетки от одного постоянного состояния к другому, которые образуют либо положительный скачок, либо контактный разрыв.

§ 12.15. Метод Годунова

В 1959 г. С. К. Годунов описал оригинальный метод решения одномерных задач со скачками, основанный на использовании уравнений Лагранжа

$$\frac{\partial}{\partial t} \mathbf{U} + \frac{\partial}{\partial x} \mathbf{F}(\mathbf{U}) = 0, \quad (12.59)$$

где

$$\mathbf{U} = \begin{bmatrix} V \\ u \\ E \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = V_0 \begin{bmatrix} -u \\ p \\ pu \end{bmatrix}, \quad (12.60)$$

как и в § 12.7. Общее в методе Годунова и в двухшаговом методе Лакса — Вендроффа состоит в том, что если $\mathbf{U}_j^n$ известны для всех j , то сначала вычисляются приближенные промежуточные значения $\mathbf{F}_{j+1/2}^{n+1/2}$, а затем вычисляются $\mathbf{U}_j^{n+1}$ по центрированной формуле

$$\mathbf{U}_j^{n+1} - \mathbf{U}_j^n = \frac{\Delta t}{\Delta x} (\mathbf{F}_{j+1/2}^{n+1/2} - \mathbf{F}_{j-1/2}^{n+1/2}). \quad (12.61)$$

Согласно второму равенству (12.60), необходимы только промежуточные значения u и p . Они получаются из физической модели, в которой распределение u , p , V , E в момент $t = n \Delta t$ аппроксимируется ступенчатыми функциями, такими, что эти величины имеют постоянные значения на каждом интервале $((j - 1/2)\Delta x, (j + 1/2)\Delta x)$. За время порядка $\Delta x/2c$ точное поведение этой модели известно: каждый разрыв немедленно распадается либо на две ударные волны, либо на две волны разрежения, либо на одну ударную волну и одну волну разрежения. (В предельных случаях одна или обе волны могут вырождаться.) Одна из этих волн движется в направлении возрастания лагранжевой координаты x , а вторая — в сторону ее убывания. Между волнами в интервале переменного x , который содержит начальный разрыв, скажем в точке $(j + 1/2)\Delta x$, u и p постоянны как по времени, так и по пространству (до тех пор, пока волны от соседних разрывов не пересекутся), в то время как V и E имеют скачки, образующие контактный разрыв в точке $(j + 1/2)\Delta x$. Постоянные значения u и p между волнами берутся в качестве $u_{j+1/2}^{n+1/2}$ и $p_{j+1/2}^{n+1/2}$.

Так как эти значения вообще не зависят от Δt , то ясно, что они не могут иметь той точности, которую имеют уравнения (12.24а, б) в двухшаговом методе Лакса — Вендроффа: ошибка аппроксимации в данном случае составляет $O(\Delta x)$, потому что величина каждого предполагаемого разрыва пропорциональна Δx , даже в области, где истинное течение является гладким. Ошибка может быть также представлена как $O(\Delta t)$, поскольку отношение $\Delta t/\Delta x$ предполагается постоянным в процессе измельчения сетки; укажем для сравнения, что ошибка аппроксимации для уравнения (12.24а) равна $O[(\Delta t)^2]$. Однако метод Годунова имеет то преимущество, что в нем любые разрывы, которые могут возникать в реальных течениях, представляются ближе к действительным, чем в методах, использующих искусственную вязкость.

Формулы для расчета гидродинамического разрыва с постоянными состояниями по обе его стороны хорошо известны; сейчас они будут получены в удобной форме при предположении, что уравнение состояния соответствует идеальному газу. Временно обозначим неизвестные $p_{j+1/2}^{n+1/2}$ и $u_{j+1/2}^{n+1/2}$ через p^* и u^* . Если волна, движущаяся вправо от точки $x = (j + 1/2)\Delta x$, представляет собой скачок, то разрыв в скорости жидкости связан с разрывом в давлении уравнением сохранения количества движения (12.35), которое в данном случае принимает вид

$$M^+ (u^* - u_{j+1}^n) = p^* - p_{j+1}^n; \quad (12.62)$$

здесь M^+ обозначает массу жидкости, проходящей через единицу поверхности скачка в единицу времени, а именно $\dot{\xi}/V_0$ ($x = \xi(t)$ — координата скачка), и определяется равенством

$$M^+ = \left(\frac{(\gamma + 1) p^* + (\gamma - 1) p_{j+1}^n}{2V_{j+1}^n} \right)^{1/2}, \quad (12.63)$$

получающимся из уравнений (12.33), (12.37) и уравнения состояния.

Если вправо движется волна разрежения, а не скачок, то в плоскости x, t центральная область, где p и u имеют постоянные значения p^* и u^* , отделяется от области, в которой они имеют постоянные значения p_{j+1}^n и u_{j+1}^n , веером, т. е. областью между двумя прямыми линиями, исходящими из точки $P_0 = ((j + 1/2)\Delta x, n\Delta t)$; между этими линиями u, V, p, ρ зависят от x и t только через переменное

$$[x - (j + 1/2)\Delta x]/(t - n\Delta t),$$

которую мы обозначим через η . Тогда второе уравнение системы (12.59) дает

$$\eta \frac{du}{d\eta} = V_0 \frac{dp}{d\eta},$$

откуда мы получаем уравнение того же вида, что и (12.62), а именно

$$M^+ (u^* - u_{j+1}^n) = p^* - p_{j+1}^n, \quad (12.64)$$

причем теперь M^+ определяется равенством

$$M^+ = \frac{1}{V_0} \frac{\int \eta du}{\int du} = \frac{1}{V_0} \bar{\eta}, \quad (12.65)$$

в котором интегрирование производится от u^* до u_{j+1}^n , т. е. M^+ есть масса, протекающая через единицу площади за единицу времени в направлении той характеристики внутри веера волны разрежения, тангенс угла наклона которой равен $\bar{\eta}$.

Чтобы получить явное выражение для M^+ , введем в первые два уравнения (12.59) переменное η и разрешим каждое из них относительно $du/d\eta$; перемножим результаты и извлечем квадратный корень:

$$du/d\eta = \sqrt{-(dV/d\eta)(dp/d\eta)}.$$

Далее, V исключается с помощью условия адиабатичности

$$\frac{V}{V_{j+1}^n} = \left(\frac{p}{p_{j+1}^n} \right)^{-1/\gamma},$$

и то приводит к уравнению

$$\frac{du}{d\eta} = \frac{2}{\gamma - 1} \frac{dc}{d\eta},$$

где через c обозначена скорость звука $\sqrt{\gamma p/V}$; следовательно, величина $2c/(\gamma - 1)$ — u остается постоянной на волне разрежения. Повторное использование условия адиабатичности приводит к нужному нам выражению

$$M^+ = \frac{\Delta p}{\Delta u} = \frac{\gamma - 1}{2 \sqrt{\gamma p_{j+1}^n V_{j+1}^n}} \frac{p_{j+1}^n - p^*}{1 - (p^*/p_{j+1}^n)^{(\gamma-1)/2\gamma}}. \quad (12.66)$$

Комбинируя формулу (12.62), которая применяется при $p^* > p_{j+1}^n$, с формулой (12.66), справедливой при $p^* \leq p_{j+1}^n$, мы находим, что

$$M^+ = \sqrt{p_{j+1}^n/V_{j+1}^n} \Phi(p^*/p_{j+1}^n), \quad (12.67)$$

где

$$\Phi(w) = \begin{cases} \sqrt{\frac{\gamma+1}{2} w + \frac{\gamma-1}{2}} & \text{при } w > 1, \\ \frac{\gamma-1}{2 \sqrt{\gamma}} \frac{1-w}{1-w^{(\gamma-1)/2\gamma}} & \text{при } w \leq 1. \end{cases} \quad (12.68)$$

Для волны, движущейся влево, аналогичным образом получаем

$$-M^-(u^* - u_j^n) = p^* - p_j^n, \quad (12.69)$$

где

$$M^- = \sqrt{p_j^n / V_j^n} \varphi(p^* / p_j^n). \quad (12.70)$$

Уравнения (12.62) и (12.69) с M^+ и M^- , выраженными через функцию $\varphi(\cdot)$, определяют p^* и u^* . Вследствие того, что p^* входит в $\varphi(\cdot)$ сложным образом, эти уравнения приходится решать итерационным методом: если приближенные значения M^+ и M^- известны (они берутся из предыдущей итерации или с предыдущего временного слоя в случае первой итерации), то приближенное значение p^* находится из уравнения

$$\frac{p^* - p_j^n}{M^-} + \frac{p^* - p_{j+1}^n}{M^+} + u_{j+1}^n - u_j^n = 0, \quad (12.71)$$

которое получается путем исключения u^* из уравнений (12.62) и (12.69); затем это значение p^* подставляется в уравнения (12.67) и (12.70) для получения следующего приближения для M^+ и M^- и т. д. После надлежащего числа итераций u^* определяется из (12.62), или из (12.69), или из уравнения

$$M^+(u^* - u_{j+1}^n) + M^-(u^* - u_j^n) = p_j^n - p_{j+1}^n, \quad (12.72)$$

получаемого исключением p^* из (12.62) и (12.69). Если эти итерации проведены до сходимости, то все три уравнения дают одно и то же значение u^* ; Годунов рекомендует использовать уравнение (12.72), вероятно потому, что оно дает лучшие результаты, когда итерации не проведены до полной сходимости.

Согласно Годунову, этот итерационный процесс сходится при условии, что ни одна из волн не является очень сильной волной разрежения. В этом случае (который имеет место весьма редко) если $p^{*(i)}$ есть результат i -й итерации, то для получения новых значений M^+ и M^- в функцию $\varphi(\cdot)$ вместо p^* подставляется выражение

$$\frac{\alpha p^{*(i-1)} + p^{*(i)}}{\alpha + 1},$$

где α — надлежащим образом выбранная положительная постоянная. Годунов рекомендует для α получаемое из предыдущей итерации значение

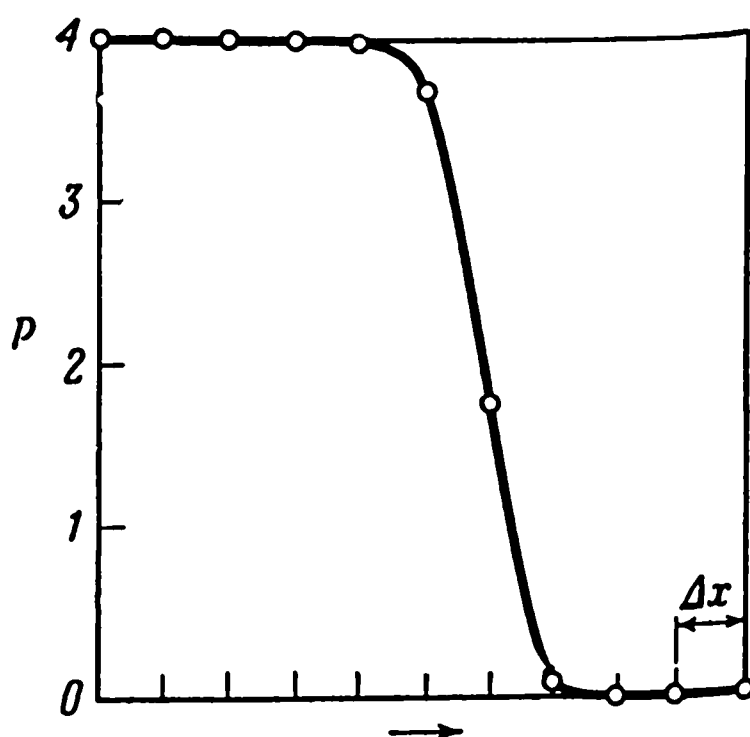
$$\alpha = \text{Max} \left(\frac{\gamma - 1}{3\gamma} \frac{1 - z}{z^{(\gamma+1)/2\gamma} - 1} - 1, 0 \right),$$

где $z = p^{*(i-1)} / (p_j^n + p_{j+1}^n)$. (Нам не ясно, почему здесь используется $p^{*(i-1)}$, а не $p^{*(i)}$.)

Если уравнение состояния не соответствует идеальному газу, то функция $\varphi(\cdot)$ задается равенством, отличным от (12.68), и параметр сходимости α также определяется другой формулой, а в остальном процедура остается той же самой.

Заметим, что в силу (12.68) функция $\varphi(\varpi)$ непрерывна в точке $\varpi = 1$ и что $\varphi(1) = \sqrt{\gamma}$; это значение соответствует распространению как слабых ударных волн, так и слабых волн разрешения со скоростью $c = \sqrt{\gamma p V}$ (т. е. со скоростью cV_0/V относительно лагранжевых координат).

На рис. 12.8 изображен профиль плоского стационарного сильного скачка, полученный этим методом. Изображенная кривая заимствована из статьи Годунова, в которой указывается,



Р и с. 12.8. Профиль давления для сильного скачка (Годунов [1959]).

что в пределах точности графического изображения профиль скачка остается неизменным для последовательных шагов вычислений. В статье указывается, что скорость согласования, к сожалению, не приводится. Этот метод широко применяется советскими специалистами.

Чтобы сравнить метод Годунова с другими методами, обсуждавшимися в этой главе, для гладкой части течения, линеаризуем уравнения (12.71) и (12.72), полагая $M^+ = M^- = c/V$ и $V = V_0$, что приводит к уравнениям

$$\begin{aligned} p^* &= p_{j+1/2}^{n+1/2} = \frac{1}{2} (p_j^n + p_{j+1}^n) - \frac{c}{2V_0} (u_{j+1}^n - u_j^n), \\ u^* &= u_{j+1/2}^{n+1/2} = \frac{1}{2} (u_j^n + u_{j+1}^n) - \frac{V_0}{2c} (p_{j+1}^n - p_j^n). \end{aligned} \quad (12.73)$$

Соответствующая линеаризация уравнения (12.59) в случае адиабатического течения дает

$$\begin{aligned} \frac{\partial u}{\partial t} &= -V_0 \frac{\partial p}{\partial x}, \\ \frac{\partial p}{\partial t} &= -\frac{c^2}{V_0} \frac{\partial u}{\partial x}; \end{aligned} \quad (12.74)$$

следовательно, (12.61) сводится к уравнениям

$$\begin{aligned} u_j^{n+1} - u_j^n &= -\frac{\Delta t}{\Delta x} V_0 (p_{j+\frac{1}{2}}^{n+\frac{1}{2}} - p_{j-\frac{1}{2}}^{n+\frac{1}{2}}), \\ p_j^{n+1} - p_j^n &= -\frac{\Delta t}{\Delta x} \frac{c^2}{V_0} (u_{j+\frac{1}{2}}^{n+\frac{1}{2}} - u_{j-\frac{1}{2}}^{n+\frac{1}{2}}); \end{aligned} \quad (12.75)$$

подставляя сюда $p_{j+\frac{1}{2}}^{n+\frac{1}{2}}$ и $u_{j+\frac{1}{2}}^{n+\frac{1}{2}}$ из (12.73), получаем окончательные уравнения

$$\begin{aligned} u_j^{n+1} - u_j^n &= -\frac{\Delta t}{\Delta x} V_0 \frac{1}{2} (p_{j+1}^n - p_{j-1}^n) + \frac{c \Delta t}{2 \Delta x} (u_{j+1}^n - 2u_j^n + u_{j-1}^n), \\ p_j^{n+1} - p_j^n &= -\frac{\Delta t}{\Delta x} \frac{c^2}{V_0} \frac{1}{2} (u_{j+1}^n - u_{j-1}^n) + \frac{c \Delta t}{2 \Delta x} (p_{j+1}^n - 2p_j^n + p_{j-1}^n). \end{aligned} \quad (12.76)$$

Эти уравнения отличаются от уравнений Лакса — Вендроффа только тем, что коэффициент при последнем члене равен $\frac{1}{2}c\Delta t/\Delta x$, а не $\frac{1}{2}(c\Delta t/\Delta x)^2$. В действительности в линеаризованном случае они эквивалентны методу Куранта, Изаксона и Риса (см. § 12.3); в этом методе уравнения записаны в характеристической форме и в каждом уравнении используются пространственные разности вперед или назад соответственно тому, положителен или отрицателен наклон характеристики. Если мы положим

$$\sigma_{\pm} = p \pm (c/V_0) u,$$

то уравнения (12.76) примут вид

$$\begin{aligned} \sigma_{+j}^{n+1} - \sigma_{+j}^n &= -\frac{c \Delta t}{\Delta x} (\sigma_{+j}^n - \sigma_{+j-1}^n), \\ \sigma_{-j}^{n+1} - \sigma_{-j}^n &= +\frac{c \Delta t}{\Delta x} (\sigma_{-j+1}^n - \sigma_{-j}^n). \end{aligned} \quad (12.77)$$

Согласно уравнению (12.74) римановы инварианты $\sigma_{\pm}$ постоянны вдоль опережающей и запаздывающей характеристик соответственно; уравнения (12.77) аппроксимируют значения $\sigma_{\pm}$ в точке сетки $(n+1, j)$ посредством проведения характеристик назад к моменту $n\Delta t$ и нахождением $\sigma_{\pm}$ в основании этой характеристики последующей линейной интерполяцией по x между двумя соседними точками сетки. Уравнения Лакса — Вендроффа для линеаризованной задачи можно интерпретировать таким же образом, за исключением того, что интерполяция в общем случае является квадратичной и использует три точки сетки $j-1, j, j+1$ в момент $n\Delta t$.

Как и для других методов в лагранжевых переменных, условие устойчивости для метода Годунова имеет вид

$$\frac{c \Delta t}{\Delta x} \frac{V_0}{V} < 1.$$

Годунов отмечает, что по крайней мере в линеаризованном виде эти уравнения обладают тем свойством, что если зависимые переменные $\sigma_{\pm}$ являются монотонными функциями от x в момент t , то вычисленные значения в момент $t + \Delta t$ также будут монотонными. Поэтому по крайней мере для слабых скачков за скачками не наблюдается колебаний, характерных для методов, использующих искусственную вязкость. Лакс и Вендрофф [1964] высказали соображения, показывающие, что для разностных уравнений повышенного порядка точности (таких, как уравнения Лакса — Вендроффа) нельзя избавиться от этих колебаний, хотя можно свести их к минимуму. Рис. 12.8 показывает, что и для сильного скачка метод Годунова дает также крутой и четкий профиль давления без последующих колебаний.

§ 12.16. Магнитная гидродинамика

В заключение этой главы мы кратко рассмотрим некоторые задачи магнитогидродинамики, описывающей течение электропроводящей жидкости в электромагнитном поле. Наиболее важной областью ее применения является физика плазмы¹⁾. Для наших целей плазму можно считать разреженным ионизованным газом, состоящим из отрицательно заряженных электронов и положительно заряженных ионов. В простейшей модели мы рассматриваем плазму как однородную жидкость с плотностью ρ , давлением p , внутренней энергией на единицу массы $\mathcal{E}$, удельным сопротивлением η и скоростью $\mathbf{u}$, несущую ток $\mathbf{J}$ в магнитном поле $\mathbf{B}$ и электрическом поле $\mathbf{E}$; электростатическими силами мы пренебрегаем. Тогда в эйлеровой форме полная система уравнений включает прежде всего законы сохранения массы, количества движения и энергии

$$\left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla\right) \rho = -\rho \nabla \cdot \mathbf{u}, \quad (12.78)$$

$$\rho \left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla\right) \mathbf{u} = -\nabla p + \mathbf{J} \times \mathbf{B}, \quad (12.79)$$

$$\rho \left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla\right) \mathcal{E} = -p \nabla \cdot \mathbf{u} + (\mathbf{E} + \mathbf{u} \times \mathbf{B}) \cdot \mathbf{J} \quad (12.80)$$

(сравни с уравнениями (12.2) — (12.4), где предполагается, что давление является скаляром и удовлетворяет уравнению состояния

$$\mathcal{E} = F(p, \rho), \quad (12.81)$$

¹⁾ В качестве введения в этот предмет см. работы Спитцера [1956] или У. Томпсона [1962].

хотя на практике часто бывает необходимо использовать давление в тензорной форме, так как магнитное поле приводит к анизотропии давления). Кроме того, мы имеем уравнения Максвелла, которые без учета тока смещения записываются так:

$$\nabla \cdot \mathbf{B} = 0, \quad (12.82)$$

$$\mu_0 \mathbf{J} = \nabla \times \mathbf{B}, \quad (12.83)$$

$$\frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = 0, \quad (12.84)$$

а также закон Ома в простейшей форме

$$\mathbf{E} + \mathbf{u} \times \mathbf{B} = \eta \mathbf{J}, \quad (12.85)$$

где μ_0 — магнитная проницаемость вакуума. Последнее уравнение используется для исключения $\mathbf{E}$ из уравнений (12.80) и (12.84); затем уравнение (12.83) может быть использовано для исключения $\mathbf{J}$, хотя часто удобнее оставить $\mathbf{J}$ как вспомогательное переменное. После этих подстановок правая часть уравнения (12.79) принимает вид

$$-\nabla \left(p + \frac{1}{2\mu_0} B^2 \right) + \frac{1}{\mu_0} (\mathbf{B} \cdot \nabla) \mathbf{B},$$

где B — модуль вектора $\mathbf{B}$, а уравнение (12.84) переходит в

$$\left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \right) \mathbf{B} = -\mathbf{B} (\nabla \cdot \mathbf{u}) + (\mathbf{B} \cdot \nabla) \mathbf{u} + (\eta/\mu_0) \nabla^2 \mathbf{B}. \quad (12.86)$$

Некоторое упрощение получается в случае симметрии слоя, так как из уравнения (12.82) тогда следует, что компонента B_x вектора $\mathbf{B}$ остается постоянной. Но для сравнения приведенных выше уравнений с гидродинамическими желательно рассмотреть сначала результат, получающийся в случае, когда $B_x = 0$ и удельное сопротивление $\eta = 0$. Тогда $\mathbf{B} \cdot \nabla = 0$ и уравнение (12.86) становится идентичным уравнению непрерывности (12.78); магнитное поле переносится вместе с жидкостью так, что $\mathbf{B}/\rho$ является функцией только лагранжевой пространственной переменной. Действительно, в этом случае полная система уравнений приводится к уравнениям гидродинамики (12.5), в которых давление заменено на $p^* = p + [1/(2\mu_0)]B^2$, а внутренняя энергия на $\mathcal{E}^* = \mathcal{E} + [1/(2\mu_0)]B^2/\rho$. В условиях устойчивости для явных разностных схем местная скорость звука теперь заменяется магнитозвуковой скоростью $(c^2 + c_A^2)^{1/2}$, где $c_A = B/(\mu_0 \rho)^{1/2}$ — скорость Альфвена.

Что касается практических задач, то основное отличие от обычной гидродинамики состоит в том, что области с нормальной плотностью плазмы и интенсивностью поля часто окружены

областями с сильными полями и очень низкими плотностями, где скорость Альфвена соответственно очень велика. В этих случаях обычно необходимо использовать неявные разностные схемы. Такие схемы с надлежащим центрированием всех величин нетрудно построить, но их решение связано с большим объемом вычислений. Так как уравнения нелинейны, то они решаются с помощью итераций и на каждой итерации необходимо решать трехдиагональную систему уравнений, аналогичную неявной системе (10.6) для волнового уравнения. Подробности читатель может найти в работе Хайна и др. [1960].

Если считать проводимость конечной, то число уравнений увеличивается, но их рассмотрение не представляло бы особых трудностей, если бы не тот факт, что в них находит отражение новое и важное физическое явление. Это диффузия магнитного поля через жидкость, описываемая последним членом уравнения (12.86). Как следствие этого конвективные члены в уравнениях должны рассматриваться весьма аккуратно, и поэтому в одномерных задачах лагранжевы переменные имеют большое преимущество. Экспериментами по численному опробованию явных двумерных схем, когда применение лагранжевых координат является затруднительным, показано, что можно использовать схему с «косой производной» (см. Робертс и Вайс [1966]). В этой схеме надлежащее центрирование достигается использованием точек сетки $(n+1, j)$ и (n, j) для образования временных разностей и $(n, j+1)$, $(n+1, j-1)$ для получения правильной компоненты пространственной разности. Эта система остается явной, если уравнения решаются в порядке возрастания j . Применительно к изэнтропическому или адиабатическому течению, описываемому при отсутствии магнитного поля уравнениями

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial X} (\rho u) = 0,$$

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial X} + (c^2/\rho) \frac{\partial \rho}{\partial X} = 0,$$

такая схема имеет вид

$$\begin{aligned} & \left(1 + \frac{\Delta t}{2 \Delta X} u_j^{n+1/2}\right) \frac{\rho_j^{n+1} - \rho_j^n}{\Delta t} + u_j^{n+1/2} \cdot \frac{\rho_{j+1}^n - \rho_{j-1}^{n+1}}{2 \Delta X} + \\ & \quad + \frac{\rho_j^{n+1} + \rho_j^n}{2} \cdot \frac{u_{j+1}^{n+1/2} - u_{j-1}^{n+1/2}}{2 \Delta X} = 0, \\ & \left(1 + \frac{\Delta t}{2 \Delta X} \cdot \frac{u_{j+1}^{n-1/2} + u_{j-1}^{n+1/2}}{2}\right) \cdot \frac{u_j^{n+1/2} - u_j^{n-1/2}}{\Delta t} + \\ & \quad + \frac{(u_{j+1}^{n-1/2})^2 - (u_{j-1}^{n+1/2})^2}{4 \Delta X} + (c^2/\rho)_j^n \frac{\rho_{j+1}^n - \rho_{j-1}^n}{2 \Delta X} = 0; \end{aligned} \tag{12.87}$$

здесь второе уравнение является линейным и явным относительно $u_j^{n+1/2}$, а ρ_j^{n+1} находится из первого уравнения. Матрица перехода для линеаризованной системы имеет характеристическое уравнение $(\lambda - \alpha)^2 + \beta^2 \lambda = 0$, где

$$\alpha = [1 + 1/2 ul (1 - e^{ik \Delta X})] / [1 + 1/2 ul (1 - e^{-ik \Delta X})],$$

$$l = \Delta t / \Delta X, \beta = cl \sin k' \Delta X / [1 + 1/2 ul (1 - e^{-ik \Delta X})].$$

Таким образом, условие фон Неймана удовлетворяется, если $|\beta| < 2|\alpha|$, т. е. если

$$l^2 (c^2 - u^2) \sin^2 k \Delta X < [2 + ul (1 - \cos k \Delta X)]^2.$$

Отсюда нетрудно видеть, что условием устойчивости является неравенство

$$(c - u) \cdot \frac{\Delta t}{2 \Delta X} < 1; \quad (12.88)$$

это согласуется с условием Куранта, Фридрихса и Леви, так как только запаздывающая характеристика рассматривается явно и пространственная сетка в этой схеме по существу удвоена.

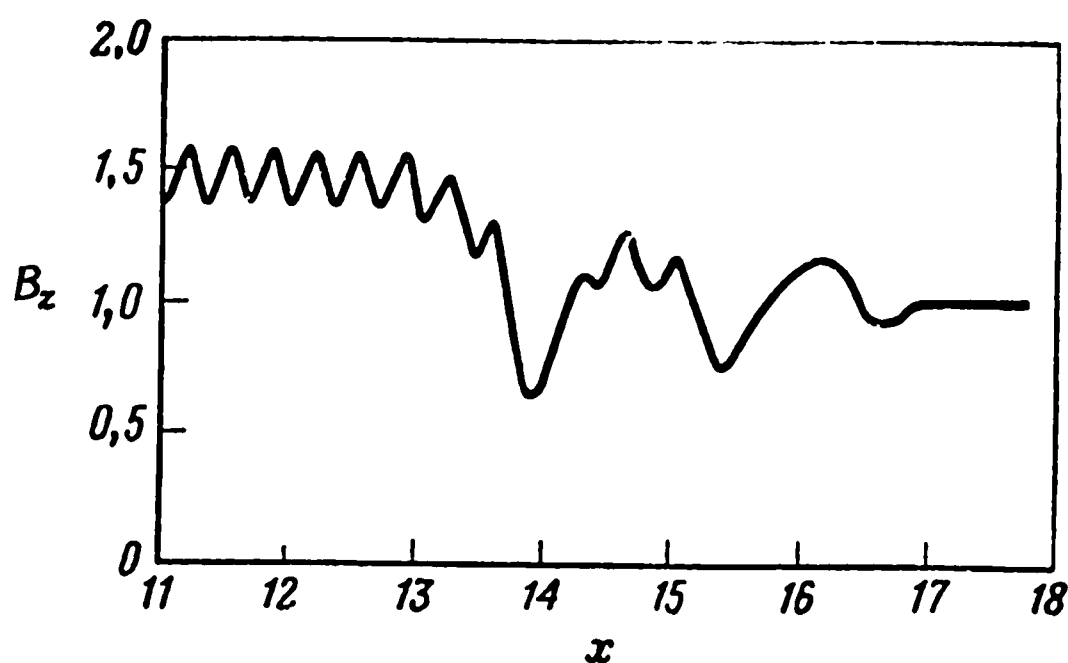
Дополнительные трудности, связанные с необходимостью аккуратного расчета диффузии магнитного поля через жидкость, возникают при наличии магнитогидродинамических скачков. Если проводимость жидкости бесконечна, то на скачке выполняются условия Гоффмана — Теллера, которые аналогичны рассмотренным выше условиям Ренкина — Гюгонио. Но обычно структура магнитного поля в области скачка очень существенна для физических приложений (в особенности в том случае, когда ограничение $\mathbf{B} \cdot \nabla \equiv 0$ ослаблено). Это требует особой точности при расчете скачка, для чего часто используется пространственная сетка, специально приспособленная к его движению.

Мы проиллюстрируем задачу о структуре скачка, рассмотрев несколько отличную физическую модель. Как было отмечено в начале этого параграфа, все введенные до сих пор уравнения основывались на предположении, что плазма является однородной жидкостью. Следствием рассмотрения плазмы, состоящей из смеси двух жидкостей, являются два уравнения движения — по одному для каждой жидкости. При сложении они дают уравнение типа (12.79), но при другой их комбинации получается уравнение для тока $\mathbf{J}$, которое заменяет введенный несколько произвольно закон Ома (12.85). Предположим, что мы объединили уравнения (12.78) — (12.84) и уравнение такого рода в систему, которая будет описывать плазму с очень малым

числом столкновений. Если мы здесь для простоты сохраним предположение, что $B_x \equiv 0$, и используем лагранжеву переменную x , введенную в § 12.4, то окончательная система уравнений примет вид

$$\begin{aligned} \frac{\partial V}{\partial t} - \frac{\partial u}{\partial x} &= 0, \\ \frac{\partial u}{\partial t} + \frac{\partial}{\partial x} \left(p + \frac{1}{2} B^2 \right) &= 0, \\ \frac{\partial \mathcal{E}}{\partial t} + p \frac{\partial V}{\partial t} &= 0, \quad \mathcal{E} = F_1(p, V), \\ \frac{\partial}{\partial t} \left[\frac{\partial^2 \mathbf{B}}{\partial x^2} - V \mathbf{B} \right] &= 0 \end{aligned} \tag{12.89}$$

(ср. с (12.8) и (12.9) и см. работу Мортонa [1964] по поводу вывода уравнений). Здесь через u обозначена компонента скорости в направлении x , отличны от нуля только компоненты



Р и с. 12.9. Типичная форма волны для «двужидкостной» плазмы при $B_x \neq 0$.

B_y и B_z вектора $\mathbf{B}$ и проведена некоторая нормировка. В силу этих уравнений поле и его первая производная непрерывны даже на скачке, а гидродинамические величины претерпевают разрыв только при достаточно большом сжатии.

Обычные конечноразностные уравнения, описанные в § 12.5, могут быть использованы для аппроксимации первых трех уравнений, а для аппроксимации последнего можно взять уравнение

$$\frac{\delta^2 B_j^{n+1}}{(\Delta x)^2} - (VB)_j^{n+1} = \frac{\delta^2 B_j^n}{(\Delta x)^2} - (VB)_j^n.$$

Изучение устойчивости в этом случае показывает, что здесь имеется много общего с уравнениями совместного распространения звука и тепла, рассмотренными в § 10.4. Энергетический

метод совершенно естественно приводит к достаточному условию устойчивости $(c^2 + c_A^2)^{1/2} (\Delta t / V \Delta x) < 1$, уже встречавшемуся в случае бесконечной проводимости. Но характеристики системы (12.89) определяются только адиабатической скоростью звука c , и метод рядов Фурье вместе с численными экспериментами показывает, что практическим условием устойчивости является неравенство

$$\left[c^2 + \left(\frac{(\Delta x)^2 V}{4 + (\Delta x)^2 V} \right) c_A^2 \right]^{1/2} \frac{\Delta t}{V \Delta x} < 1. \quad (12.90)$$

В области скачка может быть использована искусственная вязкость фон Неймана — Рихтмайера, как и для чистой гидродинамики; тогда, как и ранее, интервал устойчивости существенно сужается. Однако теперь при ее использовании необходимо соблюдать большую осторожность. Как показывает рис. 12.9, решение может иметь действительные колебания, очень похожие на те, которые были изображены на рис. 12.3 и обусловлены недостаточной искусственной вязкостью. Однако мы уже отмечали, что в области скачка решение должно находиться особенно точно, и поэтому вязкость не должна быть слишком большой.

Глава 13

МНОГОМЕРНОЕ ДВИЖЕНИЕ ЖИДКОСТИ (ГАЗА)

§ 13.1. Введение

Для краевых задач гидродинамики с двумя или более пространственными переменными применение разностных уравнений в форме Лагранжа приводит к значительной потере точности вследствие искажения лагранжевой сетки, если только не строить время от времени новую сетку, последнее же приводит к сложным и довольно неточным интерполяциям. Тем не менее были получены некоторые хорошие результаты, начиная с работы Кольски [1955], который первым подробно разработал способы составления разностных уравнений (см. также Блер и др. [1959], Гоуд [1960] и Шульц [1964b]).

Применение уравнений в форме Эйлера затруднено в тех случаях, когда имеются свободные границы или границы раздела двух сред с различными термодинамическими свойствами, так как не существует достаточно простого способа определить, какая именно среда находится в данной точке сетки в данный момент времени. Заметные успехи недавно были достигнуты (главным образом в задачах, где указанные выше трудности не возникали) благодаря использованию уравнений Лакса — Вендроффа (см., например, Бурстейн [1963, 1965]; некоторые его результаты кратко упоминаются в § 13.5).

Были развиты различные методы, использующие смешанный эйлерово-лагранжев подход; см., например, Франк и Лазарус [1964] и Нох [1964]. Эти методы часто бывают выгодными в задачах, где характер течения принципиально разный в двух различных направлениях.

Во всех этих методах для расчета скачков было введено большое количество вариантов искусственной вязкости — см. Шульц [1964a], Нох [1964] и цитированные ими работы. Двумерные варианты вязкости в методе Лакса — Вендроффа описаны в § 13.5.

Методы, использующие характеристики и бихарактеристики, разрабатывались рядом авторов, например Бруном и Хааком [1958], Ричардсоном [1964], Тэлботом [1963] и Анзорге [1963].

Де Санто и Келлер [1961] исследовали устойчивость ламинарного пограничного слоя, применив разностные методы к нестационарным уравнениям Навье — Стокса с двумя простран-

ственными переменными. Вследствие сильных сдвигов в этой задаче кажется весьма желательным применение схем, неявных относительно направления, параллельного границе. Некоторые методы решения уравнений Навье — Стокса были также предложены в работе Фромма [1964].

Метод частиц в ячейках и родственные ему методы, развитые Харлоу и его коллегами в Лос-Аламосе (Харлоу [1964]), оказались более эффективными при решении многих задач. Эти методы базируются не на разностном представлении дифференциальных уравнений с частными производными, а на некоторых простых физических моделях. Несмотря на присущую им низкую точность, даже в случае больших затрат машинного времени, эти методы оказались удивительно гибкими.

В многомерных задачах часто бывает даже неясно, являются ли физические задачи, которые нужно решать, поставленными корректно¹⁾. Линеаризованная задача о тейлоровской неустойчивости (или гельмгольцевой неустойчивости) для несжимаемой жидкости поставлена некорректно — решение не существует, если начальные данные (описывающие начальную конфигурацию поверхности раздела) не имеют бесконечное число производных, и даже если решение существует, то оно не зависит непрерывно от начальных данных.

Возможно, что эти задачи были бы корректно поставлены, если не производить линеаризацию, еще более вероятно, что они были бы корректно поставлены, если принять во внимание сжимаемость, и почти достоверно они были бы корректно поставлены, если бы учитывались также вязкость, поверхностное натяжение и теплопроводность. Но эта основная неустойчивость остается в том смысле, что малые иррегулярности могут достичь большой величины за сравнительно короткий промежуток времени. Такая неустойчивость имеет физическое происхождение, и никаким измельчением разностной сетки ее устранить нельзя. Единственный способ устранить это затруднение состоит в том, чтобы обеспечить столь гладкое представление начальных форм и течений, сколь это допускается рассматриваемой физической моделью, и не допускать, чтобы какие-либо иррегулярности вводились в последующее время.

Если, например, объектом тейлоровской неустойчивости является сферическая поверхность раздела, то совершенно неудовлетворительно представление этой поверхности некоторой аппроксимирующей псевдосферой из точек прямоугольной сетки, если эта сетка не является достаточно мелкой. Такое представление вводит иррегулярности порядка шага сетки, а это

¹⁾ Это имеет место также и для стационарных проблем — см., например, работу Моравец [1956].

может быть как раз тот порядок иррегулярности, который приводит к огромному росту иррегулярности за короткий промежуток времени в силу тейлоровской неустойчивости.

Э. Теллер (в частном сообщении) предполагает, что это затруднение может быть устранено прибавлением к уравнениям членов, представляющих влияние поверхностного натяжения на поверхности раздела. Величина этого поверхностного натяжения должна быть искусственно выбрана так, чтобы стабилизировать иррегулярность порядка шага сетки, не оказывая существенного воздействия на иррегулярности большего порядка, и должна зависеть от мгновенного ускорения поверхности раздела.

Такая схема может оказаться подходящей в случае тейлоровской неустойчивости, но гельмгольца неустойчивость может иметь место на контактном разрыве, который спонтанно порождается потоком (как это происходит при маховском отражении), и в этом случае заранее не известно, где следует вводить искусственное поверхностное натяжение.

По-видимому, мы еще далеки от обладания каким-либо универсальным методом для многомерной гидродинамики, и, вероятно, в течение некоторого времени будет примерно столько же методов, сколько проблем.

В этой главе наше внимание будет сосредоточено главным образом на развитии методов, в которых скачки и другие особенности рассматриваются более детально, чем в методах с искусственной вязкостью или в методе частиц в ячейках. Такой выбор сделан по двум причинам: во-первых, соображения экономии обычно диктуют применение в многомерных задачах весьма грубых разностных сеток, а алгоритмы, использующие искусственную вязкость, размазывают нерегулярности на такие расстояния, что детали движения большей частью утрачиваются; во-вторых, нерегулярности, которые встречаются в многомерных течениях, настолько разнообразны и сложны (некоторые из них будут обсуждены в § 13.6), что их точное численное описание требует применения алгоритмов, учитывающих специфику данной конкретной задачи.

Методы, рассматриваемые в этой главе, в настоящее время далеки от полной завершенности, и поэтому рассуждения, проводимые далее, во многих местах довольно отрывочны и условны. Тем не менее мы убеждены, что настоящие успехи всего скорее будут достигнуты именно в этом направлении.

Часто считают, что для решения многомерных задач требуются машины с большим быстродействием и большим объемом памяти, тогда как, по нашему мнению, мы чаще ограничены недостатками математических (включая численные) методов, чем недостатками машин. Нам надо больше знать о при-

роде возможных нерегулярностей, о том, как формулировать кусочно-аналитические краевые задачи в более или менее корректной форме, нам надо искать аналитические и численные методы, учитывающие эти нерегулярности, и исследовать устойчивость и точность таких методов, сочетая эмпирический и аналитический подходы.

§ 13.2. Уравнения многомерной гидродинамики

Уравнения для гладкой части течения в переменных Эйлера были приведены в виде (12.1) — (12.4). Соответствующая дивергентная форма, аналогичная уравнениям (12.16) — (12.18), для случая двух декартовых переменных X и Y имеет вид

$$\frac{\partial}{\partial t} \mathbf{U} + \frac{\partial}{\partial X} \mathbf{F}(\mathbf{U}) + \frac{\partial}{\partial Y} \mathbf{G}(\mathbf{U}) = 0, \quad (13.1)$$

где

$$\mathbf{U} = \begin{bmatrix} \rho \\ m \\ n \\ e \end{bmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{bmatrix} m \\ (m^2/\rho) + p \\ mn/\rho \\ (e + p)m/\rho \end{bmatrix}, \quad \mathbf{G}(\mathbf{U}) = \begin{bmatrix} n \\ mn/\rho \\ (n^2/\rho) + p \\ (e + p)n/\rho \end{bmatrix}; \quad (13.2)$$

здесь m и n — составляющие по осям X и Y соответственно вектора импульса, отнесенного к единице объема, e — полная энергия, отнесенная к единице объема, а p и ρ — давление и плотность. Компоненты импульса определяются через компоненты скорости жидкости как $m = \rho u$ и $n = \rho v$; если $\mathcal{E}$ — внутренняя энергия в единице объема, то полная энергия $e = \rho[\mathcal{E} + (u^2 + v^2)/2]$, а уравнение состояния принимает вид

$$p = f(\mathcal{E}, V) = f\left(\frac{e}{\rho} - \frac{m^2 + n^2}{2\rho^2}, \frac{1}{\rho}\right). \quad (13.3)$$

Фронты скачков, контактные разрывы и другие особенности обычно располагаются на поверхностях, движущихся через жидкость или вместе с ней. Проведем классификацию основных типов разрывов. Предположим, что P является точкой такой поверхности и что в окрестности этой точки поверхность имеет непрерывно меняющуюся касательную плоскость. Предположим, что ρ , p , и u имеют простой разрыв; значения этих величин в точке P по разные стороны поверхности обозначим индексами 1 и 2. Пусть λ — единичный вектор нормали к поверхности в точке P , направленный в область с индексом 2, и пусть S — скорость поверхности вдоль нормали в точке P в эйлеровой системе координат, т. е. если проведена линия через точку P

в направлении λ , то S — скорость точки пересечения поверхности с этой линией. Так как $\lambda \cdot u$ является компонентой скорости жидкости в том же самое направлении, то масса жидкости, протекающей через единичную площадку поверхности в единицу времени, равна $\rho(S - \lambda \cdot u)$; закон сохранения массы требует, чтобы

$$M = \rho_1 (S - \lambda \cdot u_1) = \rho_2 (S - \lambda \cdot u_2). \quad (13.4)$$

Изменение импульса этой массы за единицу времени равно $M(u_1 - u_2)$; сила, действующая на массу, равна $\lambda(p_1 - p_2)$, так как она рассчитана на единичную площадь поверхности. Следовательно, по второму закону Ньютона $M(u_1 - u_2) = \lambda(p_1 - p_2)$; отсюда получаются два уравнения

$$M(\lambda \cdot u_1 - \lambda \cdot u_2) = p_1 - p_2, \quad (13.5)$$

$$M(\lambda \times u_1 - \lambda \times u_2) = 0. \quad (13.6)$$

Из последнего уравнения видно, что имеется две возможности: 1) при $M \neq 0$ имеется фронт скачка, и в этом случае *тангенциальная составляющая скорости u непрерывна* при переходе через фронт, или 2) при $M = 0$ контактный разрыв, вообще говоря, сопровождается скольжением контактных поверхностей, на поверхности разрыва $p_1 = p_2$, $S = \lambda \cdot u_1 = \lambda \cdot u_2$, так что *давление и нормальная составляющая скорости u остаются непрерывными*.

При $M \neq 0$ уравнение энергии имеет вид

$$M\left(\mathcal{E}_1 + \frac{u_1^2}{2} - \mathcal{E}_2 - \frac{u_2^2}{2}\right) = p_1 \lambda \cdot u_1 - p_2 \lambda \cdot u_2, \quad (13.7)$$

где u_1 и u_2 — модули векторов u_1 и u_2 (это уравнение удовлетворяется автоматически, если $M = 0$). Условия Ренкина — Гюгонио на скачке можно вывести из соотношений (13.4) — (13.7); через удельный объем они записываются как

$$\frac{1}{V_1} (S - \lambda \cdot u_1) = \frac{1}{V_2} (S - \lambda \cdot u_2) = \sqrt{(p_1 - p_2)/(V_2 - V_1)},$$

$$\mathcal{E}_1 - \mathcal{E}_2 = (1/2)(p_1 + p_2)(V_2 - V_1), \quad (13.8a)$$

$$\lambda \times (u_1 - u_2) = 0;$$

эти уравнения отличаются от уравнений (12.37) не только числом пространственных переменных, но также и тем, что они выведены в эйлеровой системе координат, и поэтому в них вместо $d\xi/dt$ входит S . Исключая S , мы получаем равенство

$$\lambda \cdot (u_1 - u_2) = \sqrt{(p_1 - p_2)(V_2 - V_1)}, \quad (13.8b)$$

которое аналогично второму уравнению системы (12.37). Как и в случае одной пространственной переменной, отрицатель-

ные скачки ($M < 0$) исключаются по соображениям неубывания энтропии или устойчивости.

Численный расчет движения поверхности раздела в многомерном случае значительно сложнее, чем в одномерном (см. § 12.6), так как это движение обычно сопровождается скольжением жидкости вдоль поверхности, т. е. разрывностью вектора $\lambda \times u$.

Для волны разрежения не требуется граничных условий, кроме непрерывности переменных u , p , ρ . Мы не пытаемся классифицировать все другие особенности многомерных течений, но некоторые из них будут отмечены в § 13.6.

§ 13.3. Корректно и некорректно поставленные задачи

Как было определено в гл. 3, линейная задача Коши корректно поставлена, если 1) множество точных решений плотно в $\mathcal{B}$ для всех t в некотором интервале $0 \leq t \leq T$ и 2) точный разрешающий оператор ограничен для всех t из этого интервала, т. е. существует постоянная $K > 0$, такая, что любое точное решение $u(t)$ удовлетворяет неравенству $\|u(t)\| < K\|u(0)\|$ для $0 \leq t \leq T$; иначе говоря, решение непрерывно зависит от начальных данных. Для нелинейных задач условие 1) остается неизменным, но условие 2) формулируется следующим образом. Пусть $u(t)$ является точным решением задачи для начального элемента $u_0 = u(0)$; пусть $u(t) + \delta u(t)$ соответствует начальному элементу $u_0 + \delta u_0$, где δu является точным решением линеаризованной задачи (линеаризованной относительно $u(t)$ как основного решения) с начальным элементом δu_0 ; если существует постоянная $K > 0$, такая, что $\|\delta u(t)\| \leq K \|\delta u_0\|$ в некотором интервале $0 \leq t \leq T$ для всех таких $\delta u(t)$, то задача называется *корректно поставленной относительно начального элемента*. Иначе говоря, линеаризованная относительно $u(t)$ задача должна быть корректно поставленной.

В качестве первого примера некорректно поставленной задачи рассмотрим явление гельмгольцевой неустойчивости (обсуждение этого явления и родственного ему явления тейлоровской неустойчивости содержится в работе Биркгофа [1962]). Если идеальная несжимаемая жидкость имеет во всех точках выше плоскости X, Y одинаковую скорость U , направленную в положительную сторону оси X , и ту же самую скорость U в обратном направлении во всех точках ниже этой плоскости, то такое движение соответствует стационарному, но неустойчивому решению уравнений гидродинамики. Чтобы установить неустойчивость этого решения, достаточно рассмотреть в начальный момент $t = 0$ возмущение, при котором течение остается все еще безвихревым как выше, так и ниже поверхности скольжения

и асимптотически постоянным при $Z \rightarrow \pm \infty$ и в котором поверхность скольжения, заданная уравнением $z = z_0 \sin kX = Z_0 \sin(2\pi/\lambda)X$, является волнистой поверхностью с малой амплитудой Z_0 , т. е. $Z_0 \ll \lambda$.

Решение задачи с такими начальными данными показывает, что в последующие моменты времени t амплитуда колебаний $Z_0(t)$ зависит от времени как $ae^{\omega t} + be^{-\omega t}$, где a и b — постоянные, $\omega = kU$. Это решение остается справедливым до тех пор, пока $|Z_0(t)| \ll \lambda$. Выбрав λ достаточно малым (k достаточно большим), можно сделать отношение $|Z_0(t)|/|Z_0(0)|$ произвольно большим для любого фиксированного момента $t > 0$. Заметим, что при этом мы не нарушаем предположения $|Z_0(t)| \ll \lambda$, на котором основывалось приближенное решение $ae^{\omega t} + be^{-\omega t}$: мы просто выбираем k , так что $e^{\omega t} = e^{kUt}$ становится достаточно большим для данного t , а затем выбираем a таким, чтобы $ae^{\omega t} \ll 2\pi/k$ для того же t .

В этом решении отклонение скорости жидкости $\mathbf{u} = \mathbf{u}(X, Z, t)$ от основных значений $(\pm U, 0, 0)$ также пропорционально $Z_0(t)$ как выше, так и ниже поверхности скольжения. Следовательно, при любом выборе нормы для возмущения, вызванного отклонениями скорости жидкости и давления δu и δp от своих основных значений, разрешающий оператор линеаризованной задачи обязательно будет неограничен.

Задача о тейлоровской неустойчивости аналогична, за исключением того, что имеются две первоначально покоящиеся различные жидкости с плотностями ρ и ρ' выше и ниже плоскости X, Y ; если $\rho > \rho'$ и вся система имеет равномерное ускорение a в положительном направлении оси Z (или, эквивалентно, находится в гравитационном поле с ускорением a в отрицательном направлении оси Z), то малые возмущения растут аналогичным образом, но теперь степень экспоненциального роста определяется величиной

$$\omega = \sqrt{[(\rho - \rho')/(\rho + \rho')] a |k|}.$$

Здесь оператор линеаризованной задачи также неограничен для любого $t > 0$.

Если скорость жидкости и плотность имеют разрывы при $Z = 0$ и, кроме того, на поверхности раздела действует поверхностное натяжение с коэффициентом ν (здесь более подходил бы термин Биркгофа «межповерхностное натяжение»), то степень роста определяется величиной

$$\omega^2 = \frac{\rho\rho'k^2}{(\rho + \rho')^2} (U - U')^2 + \frac{\rho - \rho'}{\rho + \rho'} a |k| - \frac{\nu |k|^3}{\rho + \rho'}. \quad (13.9)$$

(см. Биркгоф, стр. 61). Если $\nu \neq 0$ (т. е. $\nu > 0$), то действие поверхностного натяжения заключается в стабилизации коротковолновых колебаний таким образом, что задача становится теперь корректно поставленной. Заметим, что стационарное решение по-прежнему неустойчиво (относительно длинных волн). Это не вызывает никаких сомнений, и, действительно, всегда можно рассчитать рост такой неустойчивости численно; особенно нежелательна неограниченная неустойчивость при $\lambda \rightarrow 0$ ($k \rightarrow \infty$) в задачах с коэффициентом $\nu = 0$.

Следствием неограниченности разрешающего оператора для задач с $\nu = 0$ является то, что если начальная форма поверхности раздела представлена рядом или интегралом Фурье, например

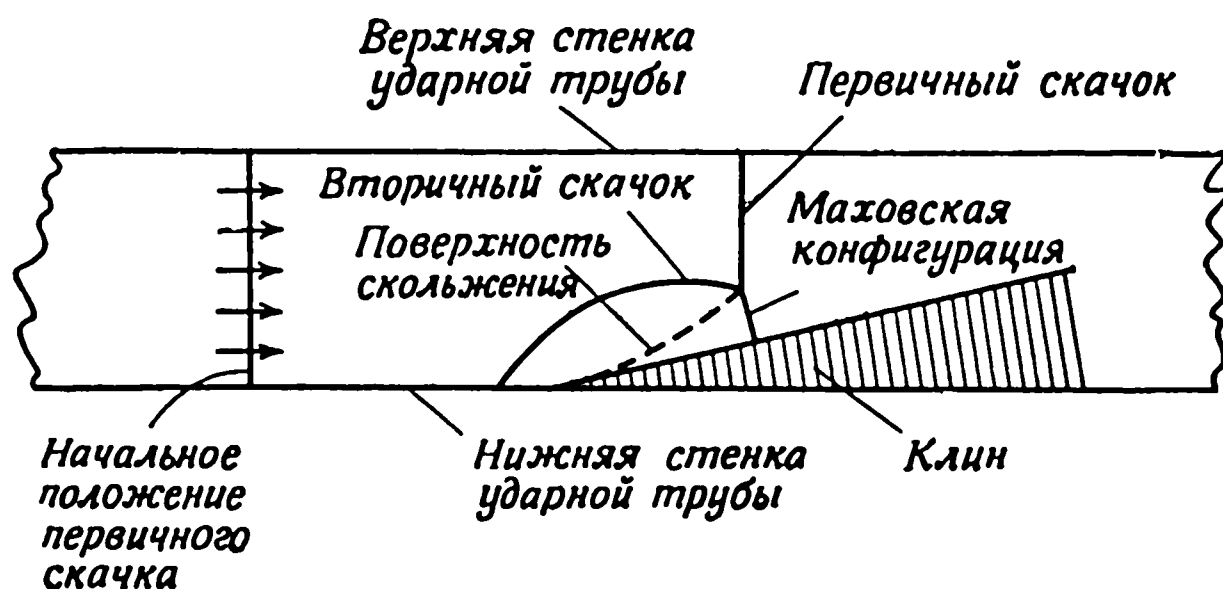
$$Z = \sum_n Z_0(n) \sin nk_0 X$$

(могут быть также члены с косинусами), то для любого $t > 0$ вообще не существует решения, если коэффициенты $Z_0(n)$ не уменьшаются быстро (в действительности по крайней мере экспоненциально) с увеличением n . Следовательно, для того чтобы существовало решение в линеаризованной теории, поверхность раздела должна быть аналитической. Мелкозернистая шероховатость, даже очень малой амплитуды, немедленно приводит к мелкомасштабному перемешиванию двух жидкостей и увеличению диффузии поверхности раздела, что наблюдалось экспериментально для $\nu = 0$.

Следует также учесть влияние сжимаемости. Из физических соображений ясно, что она будет иметь место только в том случае, когда время $\lambda/c = 2\nu/kc$, нужное для того, чтобы звуковое возмущение прошло от одного гребня волны до следующего, по меньшей мере было бы сравнимым со временем $1/\omega$, за которое амплитуда волны сможет заметно увеличиться. Из уравнения (13.9) тогда видно, что сжимаемость не может повлиять на тейлоровскую неустойчивость для коротких волн, но может оказывать некоторое влияние, хотя, вероятно, и не определяющее, на гельмгольцеву неустойчивость (в данном случае два указанных выше времени изменяются с одинаковой скоростью при $k \rightarrow \infty$). Отметим без доказательства, что если скорость скольжения $|U - U'|$ превосходит скорость звука c в определенное число раз, то гельмгольцева неустойчивость стабилизируется, тогда как для меньших скоростей скольжения задача остается некорректно поставленной при $\nu = 0$.

Если речь идет о двух различных жидкостях (например, воздухе и воде), то поверхностное натяжение, несомненно, играет известную роль и должно быть включено в формулировку задачи. Однако поверхности скольжения могут возникать и

в однородной жидкости в результате самого ее движения. Примером такого рода служит явление маховского отражения, которое показано на рис. 13.1 — см. Курант и Фридрихс [1950] и Дафф [1962]. Плоский скачок, движущийся слева направо вдоль ударной трубы, встречает клин или наклонную плоскость на нижней стенке ударной трубы. Возмущение течения начинается от острия клина в момент, когда скачок касается его, и затем распространяется дальше; границей этого возмущения служит искривленный вторичный скачок. При определенных условиях, налагаемых на угол клина и интенсивность основного



Р и с. 13.1. Маховское отражение.

скачка, первичный и вторичный скачки встречаются в некоторой точке, называемой *тройной точкой* и расположенной несколько выше клина; между этой точкой и клином два скачка сливаются в единственный (также искривленный) скачок, называемый *маховской конфигурацией*. Вниз по течению от тройной точки находится поверхность скольжения, на которой давление и нормальная компонента скорости остаются непрерывными, в то время как плотность и тангенциальная компонента скорости имеют разрывы. Элемент жидкости непосредственно ниже поверхности скольжения был сжат объединенным скачком, а элемент жидкости непосредственно выше поверхности был сжат последовательно двумя скачками до того же самого окончательного давления; сжатие, обусловленное объединенным скачком, приводит к большему росту энтропии, так как увеличение энтропии при прохождении скачка пропорционально кубу его интенсивности. Отсюда следует, что температура и удельный объем непосредственно ниже поверхности скольжения будут больше, чем непосредственно выше нее, и что для сохранения массы скорость течения вдали от скачков ниже линии скольжения также должна быть больше.

Фотографии, приведенные в статье Даффа [1962], показывают, что эта поверхность скольжения в действительности весь-

ма неустойчива относительно коротковолновых возмущений и что несколько вниз по течению она совершенно размыта. На некоторых более ранних фотографиях, например у Бликни и Тауба [1949], поверхность скольжения, выглядит совершенно гладкой. По-видимому, при некоторых условиях, когда газ впереди основного скачка находится в достаточно спокойном состоянии, любые малые возмущения, которые могут существовать вначале, не достигнут заметных размеров за то очень короткое время, в течение которого явление имеет место; однако не может быть сомнения, что неустойчивость для коротких волн действительно имеет место.

В численных расчетах нельзя полагаться на малость начальных возмущений, так как конечноразностные методы обязательно порождают возмущения, длина волны которых сравнима с шагом сетки. В § 3.9 было показано, что любая конечноразностная схема для задачи с неограниченным разрешающим оператором обязательно неустойчива. Для того чтобы расчеты некорректно поставленных задач были успешными, разностные уравнения должны иметь высокий порядок точности по пространственным переменным, начальные данные должны быть гладкими и необходимо, чтобы для всех вычислений требовалось лишь относительно небольшое число шагов по переменной t . Такие расчеты были сделаны Ван-Дайком для отсоединенного скачка. В общем случае однако нельзя надеяться на конечноразностные методы, если физическая формулировка задачи не изменена как-нибудь так, что задача становится корректно поставленной, например путем включения членов, представляющих поверхностное натяжение (на каждой поверхности скольжения или контактном разрыве) и достаточно больших, чтобы стабилизировать все возмущения, длина волны которых сравнима с шагом пространственной сетки.

§ 13.4. Двухшаговый метод Лакса — Вендроффа, или LW-метод

Для гладких течений метод Лакса — Вендроффа дает превосходные результаты. Как и в случае одной пространственной переменной (см. § 12.14), он может иногда применяться и при наличии скачков, если не требуется очень тонкое пространственное разбиение и приемлемы такие профили скачков, как показаны на рис. 12.6 и 12.7. При обобщении [Рихтмайер (1962)] двухшагового метода (12.24а) и (12.24б) на случай двух декартовых пространственных переменных X и Y удобно несколько изменить обозначения, применявшиеся в гл. 12: дробные индексы будут устранены предположением, что элементарная ячейка сетки имеет размеры $2\Delta X$, $2\Delta Y$ и $2\Delta t$ в направлениях

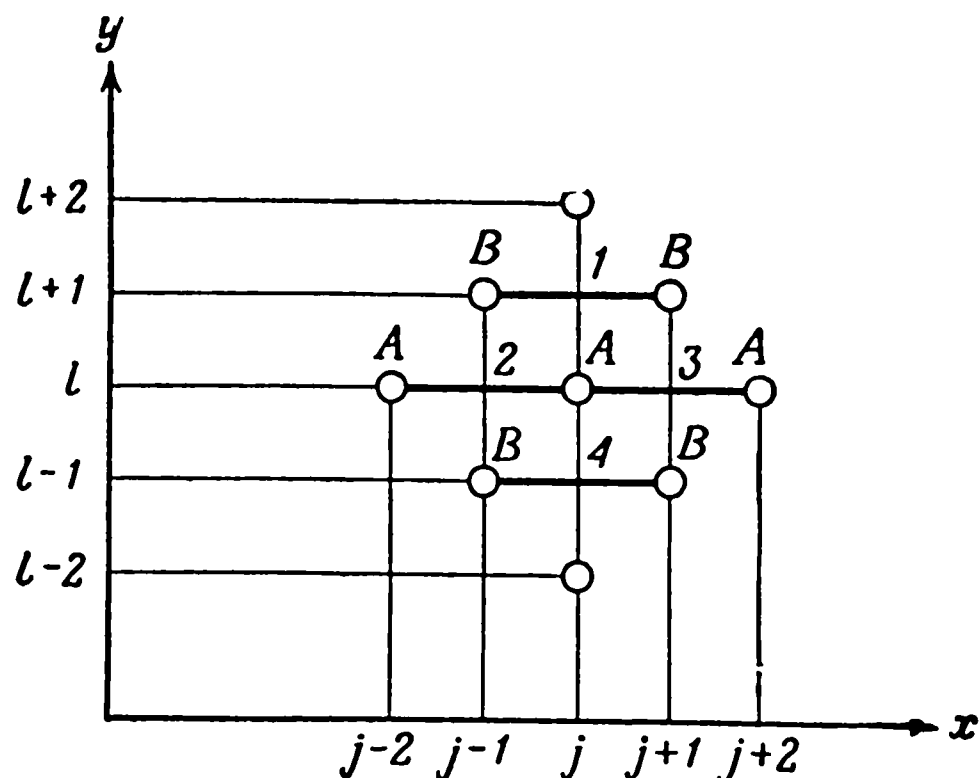
X , Y и t . Сначала находят промежуточные значения при $t = (n+1)\Delta t$ из уравнения

$$U_{jl}^{n+1} = \frac{1}{4} (U_{j+1,l}^n + U_{j-1,l}^n + U_{j,l+1}^n + U_{j,l-1}^n) - \frac{\Delta t}{2\Delta X} (F_{j+1,l}^n - F_{j-1,l}^n) - \frac{\Delta t}{2\Delta Y} (G_{j,l+1}^n - G_{j,l-1}^n), \quad (13.10)$$

где F_{jl}^n обозначает $F(U_{jl}^n)$ и т. д. и где U , $F(U)$ и $G(U)$ даны равенствами (13.2); затем окончательные значения при $t = (n+2)\Delta t$ получаются из уравнения

$$U_{jl}^{n+2} = U_{jl}^n - \frac{\Delta t}{\Delta X} (F_{j+1,l}^{n+1} - F_{j-1,l}^{n+1}) - \frac{\Delta t}{\Delta Y} (G_{j,l+1}^{n+1} - G_{j,l-1}^{n+1}). \quad (13.11)$$

Уравнения (13.10) и (13.11) не дают связи между множеством точек сетки (в пространстве — времени), где величина



Р и с. 13.2. Группа точек сетки, используемых в уравнениях (13.10) и (13.11).

$n + j + l$ четная, и множеством точек, где эта величина нечетная; поэтому половину точек сетки при желании можно опустить.

Подстановка (13.10) в (13.11) для случая $F(U) = AU$ и $G(U) = BU$, где A и B — постоянные матрицы, приводит к аналогичному (12.23) уравнению, в которое входят U^{n+2} и U^n . Это уравнение включает девятиточечную систему пространственных точек, центрированных относительно точки (j, k) , как показано на рис. 13.2; производная $\partial/\partial X$ вычислена четыре раза по парам точек, помеченных индексами 1, 2, 3 и 4, и результаты усреднены ($\partial/\partial Y$ вычисляется аналогично); $\partial^2/\partial X^2$ вычислена по трем точкам, помеченным буквой A ($\partial^2/\partial Y^2$ вычисляется аналогично); наконец $\partial^2/\partial X\partial Y$ вычислена по четырем точкам, помеченным буквой B .

Как и в случае одной пространственной переменной, эти уравнения имеют ошибку $O(\Delta^3)$, а конечноразностные аналоги

интегральных законов сохранения выполняются точно. Мы покажем, что условие устойчивости фон Неймана и предположения теоремы из § 4.11 удовлетворяются для случая $\Delta Y = \Delta X$, если

$$(|u| + c)(\Delta t / \Delta X) < 1/\sqrt{2}, \quad (13.12)$$

где c — местная адиабатическая скорость звука, а u — скорость жидкости.

Для того чтобы линеаризовать систему (13.10) и (13.11), заменим $F(U)$ и $G(U)$ на AU и BV , где матрицы $A = A(U)$ и $B = B(U)$ являются матрицами Якоби для $F(U)$ и $G(U)$, т. е. $A_{ij} = \partial F_i / \partial U_j$ и т. д. Пусть G — матрица перехода составной системы; это означает, что если в уравнениях (13.10) и (13.11) заменить U_{ji}^n на $U_0^n \exp[ik_x j \Delta x + ik_y l \Delta x]$, где U_0^n — вектор для любого n , то G будет определяться равенством $U_0^{n+2} = GU_0^n$. Мы находим, что

$$G = I - i(\cos \alpha + \cos \beta) \frac{\Delta t}{\Delta X} (A \sin \alpha + B \sin \beta) - 2 \left[\frac{\Delta t}{\Delta X} (A \sin \alpha + B \sin \beta) \right]^2, \quad (13.13)$$

где $\alpha = k_x \Delta X$ и $\beta = k_y \Delta X$.

Матрицы A и B довольно сложны, и удобнее рассмотреть вместо них упрощенные матрицы A' и B' , получающиеся из исходных уравнений в исходной эйлеровой форме (12.2) — (12.4), которые после введения скорости звука c согласно (12.7) могут быть записаны для двух переменных в следующем виде:

$$\frac{\partial U'}{\partial t} + A' \frac{\partial U'}{\partial X} + B' \frac{\partial U'}{\partial Y} = 0, \quad (13.14)$$

где

$$U' = \begin{bmatrix} \rho \\ u \\ v \\ p \end{bmatrix}, \quad A' = A'(U') = \begin{bmatrix} u & \rho & 0 & 0 \\ 0 & u & 0 & 1/\rho \\ 0 & 0 & u & 0 \\ 0 & \rho c^2 & 0 & u \end{bmatrix},$$

$$B' = B'(U') = \begin{bmatrix} v & 0 & \rho & 0 \\ 0 & v & 0 & 0 \\ 0 & 0 & v & 1/\rho \\ 0 & 0 & \rho c^2 & v \end{bmatrix}. \quad (13.15)$$

Линеаризованный вариант равенств (13.2) можно получить из (13.14) посредством преобразования подобия вида

$$dU = P dU', \quad A = PA'P^{-1}, \quad B = PB'P^{-1}. \quad (13.16)$$

Так как собственные числа матрицы перехода G инвариантны относительно преобразования подобия, условие фон Неймана для (13.2) будет тем же, что и для (13.14), и поэтому в выражении (13.13) для матрицы перехода мы можем заменить A и B на A' и B' .

Положим $C = A' \sin \alpha + B' \sin \beta$ и обозначим через θ угол между осью X и вектором, направляющие косинусы которого пропорциональны $\sin \alpha$ и $\sin \beta$, т. е.

$$\cos \theta = \frac{\sin \alpha}{\sqrt{\sin^2 \alpha + \sin^2 \beta}}.$$

Тогда

$$C = \sqrt{\sin^2 \alpha + \sin^2 \beta} \begin{bmatrix} u' & \rho \cos \theta & \rho \sin \theta & 0 \\ 0 & u' & 0 & \rho^{-1} \cos \theta \\ 0 & 0 & u' & \rho^{-1} \sin \theta \\ 0 & \rho c^2 \cos \theta & \rho c^2 \sin \theta & u' \end{bmatrix},$$

где $u' = u \cos \theta + v \sin \theta$. Собственные значения матрицы C суть

$$\lambda = \sqrt{\sin^2 \alpha + \sin^2 \beta} \begin{Bmatrix} u' \\ u' \\ u' + c \\ u' - c \end{Bmatrix},$$

и если положить $\mu = \lambda \Delta t / \Delta X$, то собственные значения матрицы G будут равны $g = 1 - i\mu(\cos \alpha + \cos \beta) - 2\mu^2$.

Следовательно,

$$\begin{aligned} |g|^2 &= (1 - 2\mu^2)^2 + \mu^2 (\cos \alpha + \cos \beta)^2 \leq \\ &\leq (1 - 2\mu^2)^2 + 2\mu^2 (\cos^2 \alpha + \cos^2 \beta) = \\ &= 1 - 2\mu^2 [\sin^2 \alpha + \sin^2 \beta - 2\mu^2] = \\ &= 1 - 2\mu^2 (\sin^2 \alpha + \sin^2 \beta) \left[1 - 2 \left(\frac{\Delta t}{\Delta X} \right)^2 \begin{Bmatrix} u' \\ u' \\ u' + c \\ u' - c \end{Bmatrix}^2 \right]. \end{aligned}$$

Наибольшее возможное значение величин в фигурных скобках составляет $|u| + c$; следовательно, при выполнении неравенства (13.12) $|g|^2 \leq 1$, что и требовалось доказать.

Согласно (13.13) G представляет собой (квадратичный) полином от матрицы $A \sin \alpha + B \sin \beta$, который равен PCP^{-1} вследствие равенств (13.16); поэтому G имеет те же самые собственные векторы, что и матрица PCP^{-1} . Матрица P легко

находится из (13.16) и

$$\det P = \rho^3 \frac{\partial}{\partial \rho} \mathcal{E}(\rho, \rho),$$

где $\mathcal{E}(\rho, \rho)$ — удельная внутренняя энергия; кроме того, P оказывается треугольной матрицей, так что $\det(P^{-1}) = (\det P)^{-1}$. Для обычных веществ при обычных условиях $d\mathcal{E}/d\rho$ ограничена снизу некоторым положительным числом — для идеального газа она равна $1/(\gamma - 1)$. Поэтому определитель матрицы, составленной из нормированных собственных векторов матрицы G , ограничен по модулю снизу некоторым положительным числом тогда и только тогда, когда то же самое справедливо для матрицы S . Нормированные собственные векторы матрицы S легко определить из вышеприведенного выражения для S , и соответствующий определитель будет равен

$$\Delta = \frac{2\rho}{c(1 + \rho^2 c^2 + \rho^2/c^2)}.$$

Несколько странный вид этого выражения объясняется тем, что U и U' являются векторами только формально и их компоненты имеют различные размерности. Мы видим, что для фиксированных ρ , ρ , u , v определитель матрицы из нормированных собственных векторов матрицы $G = G(k, \Delta t)$ не зависит от k или Δt ; следовательно, применима теорема из § 4.11, и условие фон Неймана необходимо и достаточно для устойчивости линеаризованных уравнений. Нелинейные эффекты будут обсуждаться в следующем параграфе.

Касахара [1965] рассмотрел некоторые разностные схемы, весьма похожие на схему Лакса — Вендроффа, но отличающиеся от нее тем, что некоторые величины центрированы в моменты времени $n + 1/2$ и $n - 1/2$, что кажется более естественным с точки зрения точности и упрощает некоторые из уравнений. (Идея здесь та же, что и для одномерной системы (12.10) в лагранжевых координатах, где с точки зрения центрирования разностей по времени скорости u_j^{n+1} и u_j^n в действительности относятся к моментам $n + 1/2$ и $n - 1/2$ соответственно.) Касахара получил довольно неожиданный результат, а именно некоторые схемы этого типа оказались абсолютно неустойчивыми для определенных частей течения. Это показывает, сколь осторожным надо быть при модификации стандартных разностных схем.

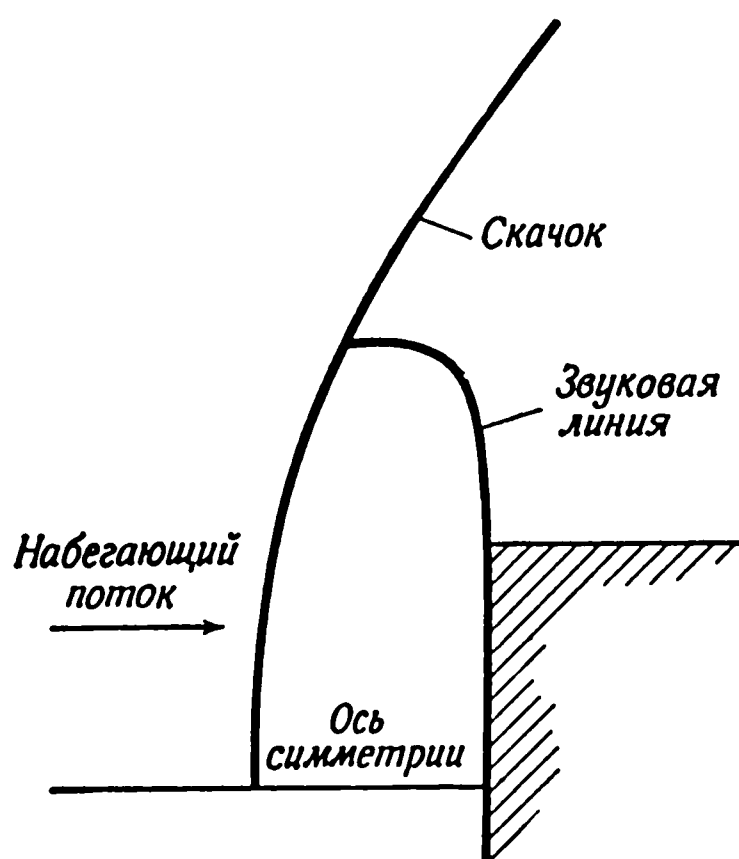
§ 13.5. Вязкость в LW-методе

Нелинейная неустойчивость в двумерных задачах еще более опасна, чем в одномерных. Согласно равенству (13.13), матрица перехода G имеет собственное значение, равное единице, если

матрица $A \sin \alpha + B \sin \beta$ имеет собственное значение, равное нулю, которое соответствует случаю $u' = 0$ или $u' = \pm c$, где u' — компонента скорости жидкости в направлении, направляющие косинусы которого пропорциональны $\sin \alpha$ и $\sin \beta$; всегда существует значение отношения $R = \sin \alpha / \sin \beta$, для которого $u' = 0$, а в области сверхзвукового течения существует два значения R , для которых $u' = \pm c$. Соответственно каждой точке X, Y, t мы имеем одну или три кривые в плоскости α, β ,

а поэтому и в плоскости k_x, k_y , вдоль которых G имеет собственное значение, равное единице; такая кривая является геометрическим местом точек удовлетворяющих уравнению $\sin \alpha / \sin \beta = \text{const}$. Уравнения нигде не являются вполне диссипативными в смысле Крайса.

Однако согласно численным экспериментам Бурстейна [1963]



Р и с. 13.3. Расчет отсоединенного скачка (Бурстейн).

для некоторых задач уравнения устойчивы и дают хорошие результаты при условии, что удовлетворяется обычный критерий устойчивости. В других задачах, согласно Бурстейну [1965] и Лapidусу (частное сообщение), неустойчивость проявляется в областях быстрого изменения (вероятно, вследствие нелинейностей или быстрого изменения коэффициентов) и быстро искажает все решение, если не ввести дополнительный член, соответствующий вязкости и аналогичный обсуждавшемуся в § 12.14.

На рис. 13.3 показаны общий характер течения между отсоединенным скачком и прямоугольным телом по расчетам Бурстейна, использовавшего метод Лакса — Вендроффа. В этой задаче существует две области очень быстрого изменения: ударный слой (который имеет в расчете конечную толщину) и окрестность угла тела, где течение изменяется очень быстро от дозвукового до сверхзвукового, и в обеих областях расчет оказывается неустойчивым, если в уравнения не включена вязкость.

Главная цель такого расчета заключается в том, чтобы найти стационарное течение путем решения краевой задачи, для которой это течение соответствует асимптотическому решению

при больших t , но проблема устойчивости здесь, вероятно, та же, что и в расчетах действительных нестационарных течений. Бурстейн обнаружил, что при введении описанной ниже вязкости известные свойства асимптотического течения передаются совершенно точно: в расчете, результаты которого представлены на рис. 13.3 (число Маха набегающего потока 4,3; идеальный газ, $\gamma = 1,4$; коэффициент вязкости $\kappa = 0,5$), поле течения после 2000 шагов было постоянным с точностью до четырех значащих цифр, отклонение плотности от постоянного значения не превышало 1%, давления — малой доли процента и т. д.

Бурстейн ввел вязкий член просто в виде суммы двух одномерных, т. е. вместо $(\Delta t/2\Delta X)\delta(Q\delta U)$, где δ — как обычно, центрированный разностный оператор, как в одномерном случае Лакса — Вендроффа, Бурстейн берет член вида

$$\frac{\Delta t}{2\Delta X} [\delta_X (Q_X \delta_X U) + \delta_Y (Q_Y \delta_Y U)], \quad (13.17)$$

где матрицы Q_X и Q_Y определены, как и в одномерном случае, но с учетом изменений только в направлениях X и Y соответственно. Это процедура оказалась удовлетворительной во всех рассматривавшихся до сих пор случаях.

Если используется вязкий член вида (13.17), то трудно вычислить собственные значения матрицы G , так как, хотя матрицы Q_X и Q_Y соответственно коммутативны с матрицами A и B , их вклад в матрицу G , главный член которого пропорционален $Q_X \sin^2 \alpha + Q_Y \sin^2 \beta$, вообще говоря не коммутативен с матрицей $A \sin \alpha + B \sin \beta$. Некоторые предварительные эксперименты Лapidуса (частное сообщение) показывают, что упрощенный вязкий член, для которого снимаются трудности исследования устойчивости, также прост в использовании и удовлетворителен на практике. Лapidус сохраняет только постоянный член g_0 квадратичного полинома $g(A)$, который входит в выражение для матрицы Q , приведенное в § 12.14; он также упрощает его путем включения только разностей скоростей u , v жидкости. Поэтому его формула совпадает с (13.17), но Q_X и Q_Y задаются равенствами

$$\begin{aligned} (Q_X)_{j+1,l} &= \kappa' |u_{j+2,l} - u_{jl}| I, \\ (Q_Y)_{j,l+1} &= \kappa' |v_{j,l+2} - v_{jl}| I. \end{aligned}$$

Тогда вклад вязких членов в матрицу перехода будет равен произведению единичной матрицы на отрицательное число; поэтому, если нуль является собственным значением матрицы $A \sin \alpha + B \sin \beta$, то соответствующее собственное значение матрицы G будет меньше единицы. Приведенное выше выражение записано в форме, удобной для двухшагового метода, в котором

мы договорились брать $2\Delta X$, $2\Delta Y$ и $2\Delta t$ в качестве основных приращений. Тогда член (13.17) заменяется выражением

$$\frac{\Delta t}{2\Delta X} [\Delta_{Ox} (Q_x \Delta_{Ox} U) + \Delta_{Oy} (Q_y \Delta_{Oy} U)]; \quad (13.18)$$

здесь для любой функции $f(X, Y)$

$$2\Delta_{Ox} f(X, Y) = f(X + \Delta X, Y) - f(X - \Delta X, Y)$$

и аналогично определяется Δ_{Oy} . В двухшаговой процедуре выражение (13.18) с U , интерпретируемым как U^n , просто добавляется к правой части уравнения (13.11).

§ 13.6. Кусочно аналитические краевые задачи

В § 13.3 было показано, что линеаризованная краевая задача о гельмгольцевой неустойчивости (а также о тейлоровской неустойчивости) не имеет решения для любого $t > 0$, если начальная поверхность раздела не является аналитической, и что, когда решение существует, оно не зависит непрерывно от начальных данных (начальной формы поверхности раздела), если отсутствует поверхностное натяжение. В общем случае для многомерных задач мы очень далеки от того, чтобы знать, существует ли решение вообще (кроме некоторых частных случаев), а если существует, то непрерывно ли оно зависит от начальных данных. Мы выскажем здесь приближенные предположения относительно существования решения. Относительно же непрерывной зависимости заметим только, что если о решении имеется достаточно сведений, чтобы в частности было известно, вероятно ли возникновение тейлоровской или гельмгольцевой неустойчивости, то можно модифицировать постановку задачи, введя поверхностное натяжение или какой-либо другой стабилизирующий механизм до начала расчета.

Существует грубое предположение о том, что если начальные данные кусочно аналитичны в пространстве, то краевая задача имеет единственное решение, которое является кусочно аналитическим в пространстве — времени по крайней мере в некотором интервале $0 \leq t \leq T$. Ясно, что мы должны определить класс краевых задач и понятие «кусочной непрерывности». Так как справедливость указанного выше предположения безусловно зависит от этих определений и так как пока мы не знаем, что надо требовать, мы намеренно избегаем полной четкости формулировок.

Класс задач, вероятно, включает в себя идеализированные задачи для сжимаемой жидкости с аналитическим уравнением состояния. Предполагается, что формулировка задачи включает дифференциальные уравнения для гладкой части течения, условия на разрывах, описанные в § 13.2, для скачков и контактных

разрывов, обычные граничные условия на внешних поверхностях и запрет отрицательных скачков. Если учитываются объемные силы, подобные силе тяжести, то они считаются заданными аналитическим безвихревым векторным полем.

Начальные данные называются *кусочно аналитическими*, если все пространство может быть разделено на клетки или области, в каждой из которых начальные функции течения являются действительными аналитическими функциями, поверхности раздела между областями являются аналитическими поверхностями и кривые, по которым эти поверхности пересекаются, представляют собой аналитические кривые.

Аналогично *решение* называется *кусочно аналитическим*, если пространство координат — времени подобным же образом делится на области, в которых параметры течения представляются аналитическими функциями. Это не означает, что разбиение пространства на области остается при $t > 0$ тем же, что и при $t = 0$; могут появиться новые поверхности раздела, как, например, когда первичный скачок достигает острия клина в задаче о маховском отражении (см. рис. 13.1) или когда внутри течения спонтанно возникает ударный фронт.

Мы имеем смутное представление о точках, которые могут оказаться критическими. Потребуем, например, что если X_0 — внутренняя точка одной из областей и если $f(X)$ — одна из функций течения, то $f(X)$ должна быть аналитической по X , т. е. существует некоторая сфера $S: \|X - X_0\| \leq a$, в которой $f(X)$ разлагается в сходящийся степенной ряд по компонентам вектора $X - X_0$; мы не говорим, должна ли быть функция $f(X)$ аналитической в точке X_0 поверхности раздела, ограничивающей эту область (в этом случае степенной ряд представлял бы $f(X)$ только в той части сферы S , которая лежит в области; в остальной части сферы S ряд представлял бы экстраполяцию $f(X)$ в одну или несколько примыкающих областей); аналогично, ничего нельзя сказать, должны ли поверхности быть аналитическими, включая границы или только до границ; наконец мы не знаем, должны ли кривые быть аналитическими, включая граничные точки или только до граничных точек.

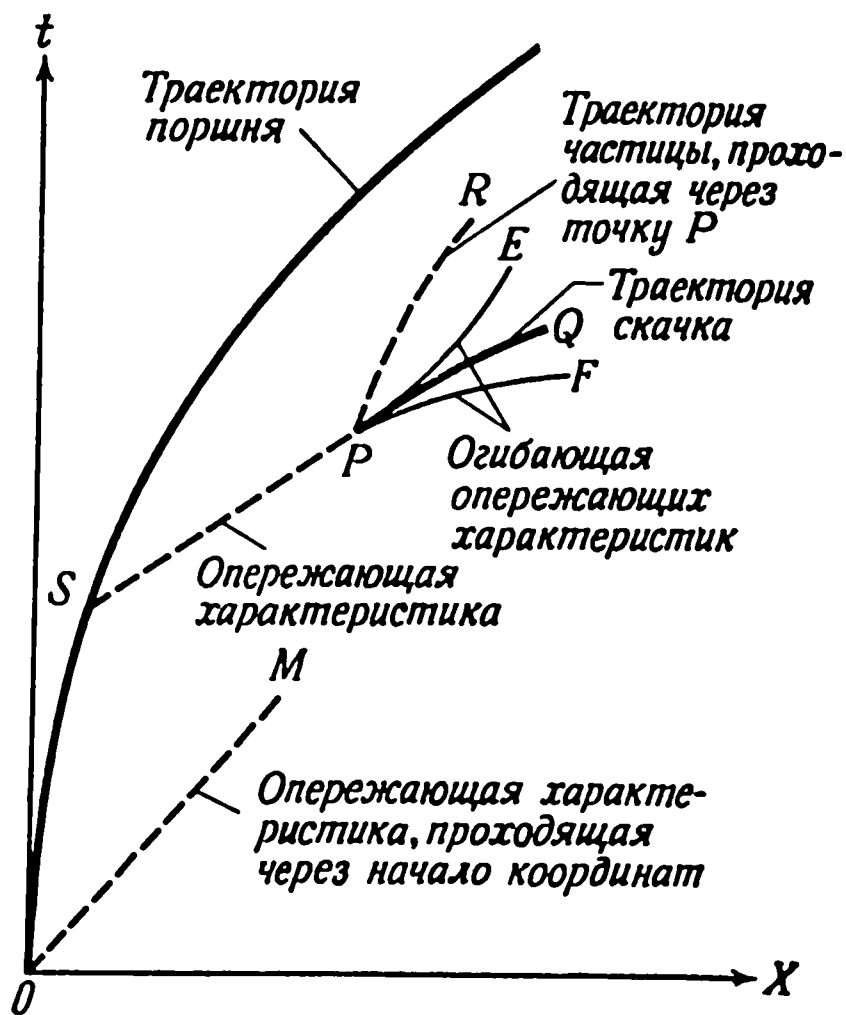
Интуиция подсказывает, что если X_0 — внутренняя точка одной из аналитических поверхностей, ограничивающих область, т. е. не располагается на ребре или в углу, то мы потребуем, чтобы в окрестности точки X_0 величины, характеризующие течение, были аналитическими функциями. Если X_0 находится на ребре, то нельзя требовать аналитичности, так как известны случаи, когда аналитичность на ребре нарушается. Известно, например, что в тройной точке в задаче о маховском отражении, т. е. в точке, где пересекаются три скачка (эта точка соответствует перпендикулярному к плоскости чертежа ребру

в действительном трехмерном течении), течение не аналитично: в окрестности тройной точки функции $X(s)$ и $Y(s)$, представляющие координаты вдоль одного из искривленных скачков (или линии скольжения), нельзя разложить в степенные ряды относительно длины дуги s , а функции течения $f(X, Y)$ нельзя разложить по степеням X и Y . Применялись разложения с дробными степенями, логарифмами и т. д., но до сих пор безуспешно; характер особенности остается неизвестным. Аналогично у острия клина при некоторых условиях осуществляется дозвуковое обтекание угла. Теория дозвукового сжимаемого неизэнтропического обтекания угла не разработана в деталях, но соответствующая особенность, вероятно, может быть описана разложением по пространственным переменным с дробными степенями. В этих примерах основные переменные течения при приближении к особенности остаются ограниченными и, возможно, это должно быть общим требованием. Пример веера характеристик в центрированной волне разрежения показывает, что при приближении к особенности основные величины течения не всегда имеют единственные предельные значения.

Для задач, включающих зависимость от одной пространственной переменной, некоторые предположения указанного выше типа были сформулированы, например, Гёльднером [1949], Рождественским [1960] и Лоджмэном [1965]. Лоджмэн рассмотрел гиперболическую систему второго порядка, аналогичную системе (третьего порядка) уравнений гидродинамики, вместе с условиями Ренкина — Гюгонио и условием для энтропии; в его задаче зависимые переменные $u(x, t)$, $v(x, t)$ были достаточно гладкими функциями x при $t > 0$, за исключение точки $x = x_0$, где они имели простой разрыв. Соответствующая задача, в которой $u(x, 0)$ и $v(x, 0)$ являются постоянными, за исключением точки $x = x_0$, хорошо известна; решение содержит или два скачка, или две центрированных волны разрежения, или скачок и центрированную волну разрежения, начинающиеся от $x = x_0$ и движущиеся по жидкости в противоположных направлениях; между этими волнами скорости u и v постоянны. Лоджмэн доказал, что подобное решение, вообще говоря, существует и единственно при надлежащих условиях; фронт скачка и голова и хвост волны разрежения в общем случае движутся с различными скоростями, а u и v между волнами не остаются постоянными, но являются гладкими функциями. Для многомерных задач, по-видимому, нельзя рассчитывать на то, что существенные успехи будут достигнуты в ближайшее время.

Аналогичные предположения можно сделать для других классов задач, например для уравнений Навье — Стокса, для задач магнитной гидродинамики и гидродинамики с переносом тепла или химическими реакциями.

В качестве частного случая рассмотрим спонтанное возникновение скачка. Если поршень, движущийся со скоростью, которая непрерывно возрастает от нуля, вдвигается в трубу, заполненную первоначально покоящимся газом с давлением $p > 0$, то возникает скачок, как это схематически показано на рис. 13.4. Интенсивность этого скачка в точке P равна нулю. При определенных предположениях относительно движения поршня новых скачков не возникнет. Если уравнение состояния является соотношением между давлением p и плотностью ρ , не зависящим от энтропии, а траектория поршня $X(t)$ при $t > 0$ представляется аналитической функцией, то траектория скачка (кривая PQ) является единственной особенностью течения, за исключением особенности высшего порядка на той опережающей характеристике OM , которая проходит через точку, соот-



Р и с. 13.4. Спонтанное возникновение скачка.

ветствующую началу движения поршня. В этом случае все решение задачи может быть проведено методом характеристик, как это сделано, например, Курантом и Фридрихсом [1950]. Любая основная функция $f(X, t)$ течения является аналитической выше кривой PQ , на этой кривой и на обеих ее сторонах, за исключением самой точки P , где имеется особенность. Поверхность, представляющая $f(X, t)$ в пространстве (X, t, f) , образует складки таким образом, что f формально является трехзначной в области между огибающими PE и PF ; вдоль траектории скачка PQ , расположенной между PE и PF , происходит разрывной переход с верхней части этой поверхности (которая дает правильные значения f слева от PQ) на нижнюю ее часть (которая дает значения f справа от PQ). Для более общего уравнения состояния существует дополнительная особенность на траектории частицы PR , где третья производная от энтропии по X становится бесконечной.

В качестве второго частного случая рассмотрим гельмгольцеву неустойчивость на тупом клине, подобно тому, что показано на рис. 13.5 в точке P на начальной поверхности раздела.

Как было указано в § 13.3, попытка найти решение путем линеаризации (которая включает предположение, что поверхность раздела близка к плоской) и методом рядов Фурье полностью несостоятельна, если начальная поверхность раздела не является аналитической. Предположим тем не менее, что некоторое аналитическое (или кусочно аналитическое) решение существует в некотором интервале $(0, t_0)$ для идеализированной задачи, в которой начальная поверхность является аналитической всюду, за исключением ребра клина. (Неточности, вероятно, должны привести к мелкомасштабному перемешиванию, как это описывалось выше в конце § 13.3.)



Р и с. 13.5. Предполагаемая конфигурация поверхности раздела при гельмгольцевой неустойчивости.

Идеализируем задачу далее, пренебрегая сжимаемостью, вязкостью и поверхностным натяжением и предполагая, что при $t = 0$ поле скоростей задается известным решением для несжимаемого течения около угла твердого тела; в координатах X, Y , показанных на рис. 13.5, потенциал начальной скорости представляет собой действительную часть функции $Z^{\pi/\beta}$, где $Z = X + iY$, а β — угол клина, как показано на рисунке. Для $t > 0$ давление на поверхности раздела считается равным нулю, т. е. она является свободной поверхностью. Предполагается, что движение жидкости выше поверхности раздела описывается уравнениями динамики для несжимаемой жидкости. Существует также граничное условие на поверхности раздела, которое гласит, что нормальная компонента скорости жидкости должна равняться нормальной компоненте скорости самой поверхности раздела. Наконец, требуется, чтобы на больших расстояниях от особенности потенциал скорости совпадал асимптотически со своим начальным значением, а именно, действительной частью $Z^{\pi/\beta}$.

Задача, сформулированная таким образом, не содержит масштаба длины; если решение существует и единственно, оно должно быть автомодельным, т. е. потенциал скорости $\varphi(X, Y, t)$ должен иметь вид $g(t)\Phi(\xi, \eta)$, где $\xi = X/f(t)$ и $\eta = Y/f(t)$, и аналогично в момент t радиус-вектор $Z = X + iY = Z(s, t)$ частицы, которая в начальный момент находилась на поверхности раздела на расстоянии s по дуге от угла, должен иметь вид $Z(s, t) = f(t)\zeta[s/f(t)]$. Мгновенные фотографии поверхности раздела, сделанные в моменты t_1 и t_2 , окажутся идентичными с точностью до увеличения. Если эти выражения подставить в дифференциальные уравнения и в граничные условия, то окажется, что задача, действительно, сведется к краевой задаче, содержащей функции $\Phi(\xi, \eta)$ и $\zeta(\sigma)$, при условии что $f(t) = t^\alpha$, где $\alpha = (2 - \pi/\beta)^{-1}$ и $g(t) = t^\gamma$, а $\gamma = \pi\alpha/\beta$.

Можно предположить, что поверхность раздела окажется примерно такой, как показано на рис. 13.5. Так как $0 < \alpha < 1$ для $\beta > \pi$, кривая, представляющая поверхность раздела, начинает распространяться от ребра клина в момент $t = 0$ с бесконечной начальной скоростью. Кажется правдоподобным, что эта кривая имеет колебания относительно начальных прямых линий с быстро уменьшающейся амплитудой (и, возможно, увеличивающейся длиной волны) по мере удаления от начала координат. Однако вопрос существования и единственности решения этой упрощенной задачи остается в настоящее время открытым.

Можно представить себе и другие конфигурации, в которых, например, свободная поверхность имеет точку возврата и, возможно, вихревой слой, начинающийся от этой точки. Однако в любом случае изменение конфигурации, вероятно, должно подчиняться принципу подобия.

§ 13.7. Программа развития методов

Справедливость предположений, описанных в предыдущем параграфе, имела бы значение не только для установления существования и единственности решений, но и для разработки программы развития численных методов, которую мы здесь кратко наметим.

В гладкой части течения (в области аналитичности) можно применять ортогональную сетку в пространстве координат — времени и аппроксимировать дифференциальные уравнения разностными уравнениями, которые в принципе могут иметь произвольно высокий порядок точности, так как функции аналитичны, но на практике обычно используется двухшаговый метод Лакса — Вендроффа или ему подобный. Разрывы и особенности всех типов в той степени, насколько заранее известна их при-

рода, могут быть исследованы специальными методами, подобными процедурам подгонки скачков, чтобы объединить результаты разностных расчетов в различных областях и удовлетворить условиям на разрыве с точностью, соответствующей точности расчета гладкого течения.

Если N (равное 2, 3 или 4) — размерность пространства координат — времени, то особенностями, соответствующими гиперповерхностям размерности $N - 1$, могут быть: 1) ударные волны, 2) контактные разрывы (которые могут быть также поверхностями скольжения), 3) поверхности, на которых основные величины течения непрерывны, но их производные некоторого порядка разрывны. Мы придерживаемся точки зрения, что не стоит разрабатывать специальные методы для особенностей типа 3), по крайней мере в настоящее время (хотя сделать это было бы легче, чем осуществить подгонку скачка); вместо этого мы предположим, что для такой гиперповерхности расчет проводится на прямоугольной сетке так, как будто никакой особенности не существует. В § 13.9 будет описан упрощенный случай подгонки скачков и приведены (весьма кратко) сведения о более общем случае подгонки скачка и методах расчета контактных разрывов.

Примерами особенностей, соответствующих гиперповерхностям размерности $N - 2$, являются: для $N = 2$ (одна пространственная переменная) 1) столкновение двух параллельных плоских или концентрических сферических скачков, 2) мгновенное совпадение скачка с контактным разрывом или поверхностью раздела двух сред, 3) столкновение двух свободных поверхностей, 4) фокусировка сходящегося сферического скачка в центре симметрии. Для $N = 3$ примерами являются: 5) тройная точка, где в случае маховского отражения встречаются три скачка, 6) пересечение двух некомпланарных скачков. В примерах 1) — 4) характер особенности известен, и для того чтобы найти специальные численные процедуры подгонки, можно использовать степенные ряды или другие разложения; мы не будем описывать эти процедуры, так как они обычно имеют весьма частный вид применительно к конкретным задачам. В примерах 5) и 6) характер особенности не известен.

§ 13.8. Характеристики в двумерных течениях

При обсуждении подгонки скачка в одномерном случае в § 12.9 было указано, что информация о течении в некотором смысле переносится вдоль характеристик или характеристических кривых в плоскости X, t и что только одна из трех характеристик, проходящих через данную точку, имеет такое направление, что способна переносить информацию к скачку из тече-

ния позади него. Эти соображения будут теперь обобщены на двумерное течение.

Одномерные уравнения в характеристической форме были даны в гл. 12 — уравнения (12.30). В каждом из этих трех уравнений все зависимые величины дифференцируются в одном и том же направлении в плоскости X, t . В двумерном течении существует только одна линейная комбинация первоначальных дифференциальных уравнений, в которой все величины оказываются продифференцированными в одном направлении; это уравнение $D\mathcal{E} + pD(1/\rho) = 0$, где D — оператор дифференцирования вдоль траектории частицы, которое означает, что энтропия постоянна вдоль траектории частицы в гладкой части течения. Даже это уравнение является особенностью гидродинамических уравнений; самое большее, на что можно надеяться для общей гиперболической системы с тремя независимыми переменными, это найти такую линейную комбинацию исходных уравнений, что зависимые величины оказываются продифференцированными в компланарных направлениях. Такие уравнения называются уравнениями в *характеристической форме* для двумерного течения, а соответствующая плоскость называется *характеристической плоскостью*. Далее будет показано, что существует бесконечно много характеристических плоскостей, проходящих через данную точку (X_0, Y_0, t_0) , а именно, любая плоскость, касательная к траектории частицы в точке (X_0, Y_0, t_0) , является характеристической и любая плоскость, касательная в точке (X_0, Y_0, t_0) к звуковому конусу с вершиной в (X_0, Y_0, t_0) , также является характеристической.

Чтобы найти возможные ориентации характеристической плоскости, нужно четыре уравнения (12.2) — (12.4) (второе уравнение в действительности представляет два уравнения, так как скорость теперь имеет две компоненты) умножить на коэффициенты $\alpha, \beta, \gamma, \delta$ соответственно и сложить результаты, а внутреннюю энергию $\mathcal{E}$ выразить через p и ρ с помощью уравнения состояния. Окончательное уравнение принимает вид

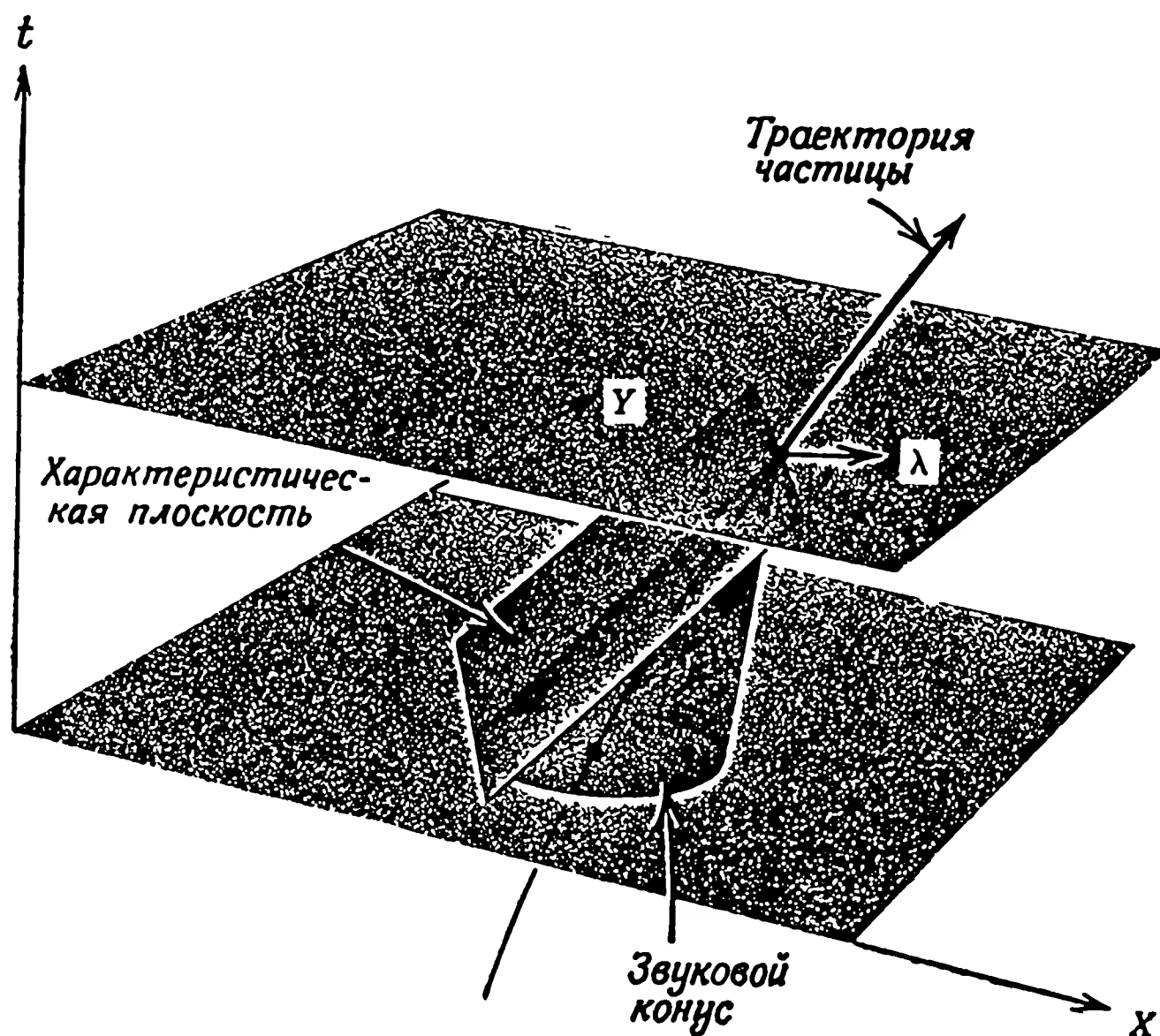
$$\eta_1 \nabla p + \eta_2 \nabla \rho + \eta_3 \nabla u + \eta_4 \nabla v = 0, \quad (13.19)$$

где ∇ — векторный оператор, компоненты которого равны $(\partial/\partial X, \partial/\partial Y, \partial/\partial t)$. Функции η_i линейны относительно $\alpha, \beta, \gamma, \delta$ и зависят также от p, ρ, u, v , а следовательно, от X, Y, t . Требование, что η_i компланарны, означает существование такого ненулевого вектора, что

$$\xi \cdot \eta_i = 0 \quad (i = 1, 2, 3, 4); \quad (13.20)$$

вектор ξ , таким образом, является нормалью к характеристической плоскости. При фиксированных X, Y, t это дает четыре линейных уравнения с четырьмя неизвестными $\alpha, \beta, \gamma, \delta$; они

имеют ненулевое решение только тогда, когда $\Delta = 0$, где $\Delta = \Delta(\xi)$ является определителем системы и представляет собой полином самое большее четвертой степени от трех компонент вектора ξ . Так как длина вектора произвольна, уравнение $\Delta = 0$ определяет однопараметрическое семейство возможных



Р и с. 13.6. Пространственно-временная диаграмма, показывающая траекторию частицы и звуковой конус.

ориентаций характеристической плоскости, проходящей через данную точку X, Y, t .

Уравнения в характеристической форме, полученные описанным выше способом, имеют вид

$$\frac{Dp}{Dt} - c^2 \frac{D\rho}{Dt} = 0, \quad (13.21)$$

$$\mu \cdot \left(\frac{D\mathbf{u}}{Dt} + \frac{1}{\rho} \nabla p \right) = 0, \quad (13.22)$$

$$\rho c \left(\lambda \frac{D}{Dt} + c \nabla \right) \cdot \mathbf{u} + \left(\frac{D}{Dt} + c \lambda \cdot \nabla \right) p = 0; \quad (13.23)$$

все векторы, входящие в эти уравнения, являются двумерными (в плоскости X, Y), в частности $\mathbf{u} = (u, v)$, а $\nabla = (\partial/\partial X, \partial/\partial Y)$; кроме того, c — адиабатическая скорость звука (функция от p и ρ), а λ и μ — произвольные единичные векторы в плоско-

сти X, Y ; наконец, D/Dt означает оператор

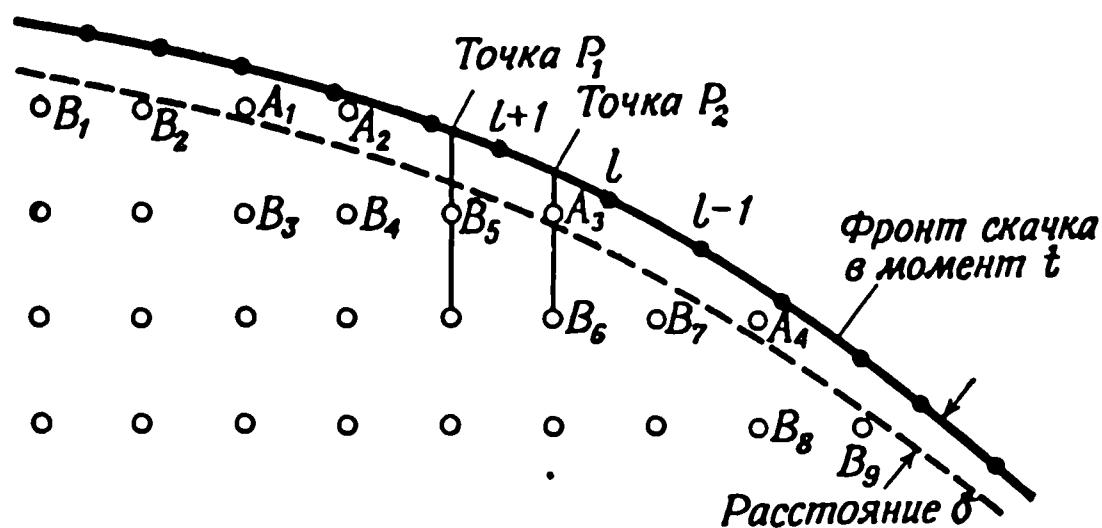
$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla = \frac{\partial}{\partial t} + u \frac{\partial}{\partial X} + v \frac{\partial}{\partial Y} \quad (13.24)$$

дифференцирования вдоль траекторий частиц (мировых линий). В уравнении (13.22) направления дифференцирования находятся в плоскости, касательной к траектории частицы и параллельной μ ; в (13.23) они находятся в плоскости, касательной к звуковому конусу и такой, что прямая, по которой эта плоскость пересекается с плоскостью X, Y , перпендикулярна к λ (см. рис. 13.6).

С этими уравнениями и соответствующими характеристическими плоскостями мы будем иметь дело в следующем параграфе.

§ 13.9. Подгонка скачков в двумерных задачах

Чтобы рассмотреть сначала простейший случай, предположим, что скачок движется в первоначально невозмущенной среде, так что впереди него давление, плотность и две компоненты скорости жидкости имеют известные постоянные значения p_2, ρ_2, u_2 и v_2 (в обозначениях § 13.2), тогда как позади него



Р и с. 13.7. Расположение точек сетки и точек скачка.

они являются функциями X, Y, t ; здесь, как всегда, X и Y — это декартовы эйлеровы координаты. В любой момент $t = n\Delta t$ фронт скачка располагается на некоторой кривой в плоскости X, Y ; в численном расчете эта кривая представлена последовательностью точек с координатами $(X, Y) = (\xi_l^n, \eta_l^n) = \xi_l^n$ вдоль нее, $l = 1, 2, \dots$, как показано на рис. 13.7, где показана и прямоугольная сетка, узлы которой изображены светлыми кружками. Величины p, ρ, u, v в точке ξ_l^n в момент $n\Delta t$ обозначены через $p_l^n, \rho_l^n, u_l^n, v_l^n$ (это значения непосредственно позади скачка; значения на фронте — p_2, ρ_2, u_2, v_2), тогда как величины в узлах основной сетки обозначены через p_{jk}^n и т. д.,

как в § 13.4. Единичный вектор в плоскости X, Y , перпендикулярный скачку и направленный вперед, обозначен через λ_l^n .

Предположим, что в момент $t = n\Delta t$ все величины в точках скачка и узлах сетки позади скачка известны и что требуется найти эти величины в момент $t = (n+1)\Delta t$. Нужно найти и новое положение $(\xi_l^{n+1}, \eta_l^{n+1}) = \xi_l^{n+1}$ каждой точки скачка; это положение отчасти произвольно, поскольку оно находится на фронте скачка в момент $t = (n+1)\Delta t$ где-то между точками ξ_{l+1}^{n+1} и ξ_{l-1}^{n+1} и довольно близко к старому положению ξ_l^n , так что единственная компонента скорости $\dot{\xi}_l$, которая имеет физический смысл, есть компонента в направлении λ_l . Для простоты мы поэтому положим $\dot{\xi}_l = S_l \lambda_l$, где S_l — скорость скачка, входящая в уравнения (13.8а); в разностной форме это можно записать как

$$\frac{1}{\Delta t} (\xi_l^{n+1} - \xi_l^n) = \frac{1}{2} (S_l^{n+1} \lambda_l^{n+1} + S_l^n \lambda_l^n). \quad (13.25)$$

Отметим здесь, что в целом процедура подгонки скачков является существенно неявной. Две компоненты (13.25) дают два уравнения с пятью неизвестными; вместе с другими уравнениями, приведенными ниже, они образуют сложную систему нелинейных уравнений, которая на практике должна решаться итерационным путем.

Из условий Ренкино — Гюгонио (13.8а) и (13.8б) находим, что

$$S_l^{n+1} - \lambda_l^{n+1} \cdot \mathbf{u}_2 = \frac{1}{\rho_2} \sqrt{\frac{p_l^{n+1} - \bar{p}_2}{1/\rho_2 - 1/\rho_l^{n+1}}}, \quad (13.26)$$

$$\lambda_l^{n+1} \cdot (\mathbf{u}_l^{n+1} - \mathbf{u}_2) = \sqrt{(p_l^{n+1} - p_2)(1/\rho_2 - 1/\rho_l^{n+1})}, \quad (13.27)$$

$$F(p_l^{n+1}, \rho_l^{n+1}) - F(p_2, \rho_2) = \frac{1}{2} (p_l^{n+1} + p_2)(1/\rho_2 - 1/\rho_l^{n+1}), \quad (13.28)$$

$$\lambda_l^{n+1} \times (\mathbf{u}_l^{n+1} - \mathbf{u}_2) = 0, \quad (13.29)$$

где $F(p, \rho)$ равно $\mathcal{E}$ согласно уравнению состояния (12.1). Неизвестные в этих уравнениях относятся к моменту $t = (n+1)\Delta t$. Единичный вектор λ_l^{n+1} может быть выражен через положение точки скачка l и двух ее соседних точек путем предварительного вычисления вектора λ^* , касательного к кривой в точке ξ_l^{n+1} :

$$\lambda^* = \alpha (\xi_{l+1}^{n+1} - \xi_l^{n+1}) + \beta (\xi_l^{n+1} - \xi_{l-1}^{n+1}), \quad (13.30)$$

где α и β — коэффициенты (функции от ξ), выбранные так, чтобы λ^* был единичным вектором и точки, соседние с точкой

ξ_l^{n+1} , входили с весами, обратно пропорциональными их расстояниям от нее. (Это эквивалентно построению некоторой кривой второго порядка по трем точкам и определению λ^* как касательной к этой кривой в точке ξ_l^{n+1} .) Тогда вектор λ , определяемый равенством

$$(\lambda_x, \lambda_y) = (-\lambda_y^*, \lambda_x^*), \quad (13.31)$$

является искомой единичной нормалью λ_l^{n+1} ; если индекс l возрастает в направлении вдоль кривой так, что при этом сжатая область остается слева, то λ_l^{n+1} является нормалью, направленной из этой области наружу.

Учитывая векторные уравнения (13.25), (13.30) и (13.31) дважды и промежуточные уравнения один раз, мы получаем таким образом 10 уравнений с 11 неизвестными, а именно S , p , ρ и компонентами векторов u , ξ , λ , λ^* в точке l скачка в момент $(n+1)\Delta t$. (Заметим в этой связи, что (13.29) рассматривается как одно уравнение, так как только его Z -компонента несет информацию; это уравнение может быть записано в виде $\lambda^* \cdot (u_l^{n+1} - u_2) = 0$.) Как и при подгонке одномерных скачков, необходимое добавочное уравнение получается из условия, что при надлежащей разностной аппроксимации одно из дифференциальных уравнений должно удовлетворяться не только в точках основной сетки, но и в каждой точке скачка. Необходимо решить: 1) какое дифференциальное уравнение использовать и 2) какую разностную аппроксимацию выбрать для этого уравнения. Кроме того, необходимо решить, 3) как получить новые значения p_{jk}^{n+1} и т. д. в тех точках сетки, которые так близки к скачку, что обычные разностные уравнения (например, Лакса — Вендроффа) для них непригодны. (На практике давление p_{jk}^{n+1} в этих точках, а также и регулярных точках сетки рассчитывается перед подгонкой скачка.) Ясно, что на все эти вопросы можно дать самые различные ответы; ответы, данные ниже, приводятся главным образом в качестве иллюстрации, так как относительные преимущества многих возможных процедур полностью еще не исследованы и по всей вероятности ни одна из этих процедур не окажется пригодной для всех задач.

Как и при подгонке одномерного скачка, если дифференциальное уравнение, выбранное для дополнения условий на разрыве, как было описано выше, является одним из уравнений в характеристической форме (13.21), (13.22) или (13.23), то простые физические соображения подсказывают однозначный выбор этого уравнения. В одномерном течении информация в известном смысле распространяется по характеристическим направлениям в плоскости X, t ; в том же смысле в двумерных задачах

информация переносится вдоль характеристических плоскостей. Так как нам нужна информация из потока позади скачка, мы выбираем характеристическую плоскость, проходящую через точку $(\xi_l^{n+1}, (n+1)\Delta t)$ в пространстве координат — времени так, что ее «прошлая» часть лежит, насколько это возможно, в области позади скачка, и этот выбор является однозначным: мы должны выбрать плоскость, касательную к звуковому конусу, пересечение которой с плоскостью $t = (n+1)\Delta t$ является касательной к фронту скачка и которая, если ось t предполагается вертикальной, наклонена вперед максимально в направлении скачка. Иначе говоря, мы должны выбрать уравнение (13.23) с вектором λ , направленным по нормали скачку, именно, $\lambda = \lambda^{n+1}$.

Для простоты мы выпишем только разностную аппроксимацию первого порядка для уравнения (13.23), соответствующую замене уравнения (13.25) простейшим уравнением $\xi_l^{n+1} - \xi_l^n = S_l^n \lambda_l^n \Delta t$. Дивергенция скорости может быть записана как $\nabla \cdot \mathbf{u} = (\lambda \cdot \nabla)(\lambda \cdot \mathbf{u}) + (\lambda^* \cdot \nabla)(\lambda^* \cdot \mathbf{u})$, где λ^* — единичный вектор (в плоскости X, Y), ортогональный λ .

Тогда (13.23) можно записать как

$$\frac{\delta}{\delta t} p + \rho c \frac{\delta}{\delta t} (\lambda \cdot \mathbf{u}) + (\lambda^* \cdot \nabla)(\lambda^* \cdot \mathbf{u}) = 0,$$

где

$$\frac{\delta}{\delta t} = \frac{\partial}{\partial t} + (\mathbf{u} + c\lambda) \cdot \nabla$$

— оператор дифференцирования по времени вдоль линии касания характеристической плоскости и звукового конуса.

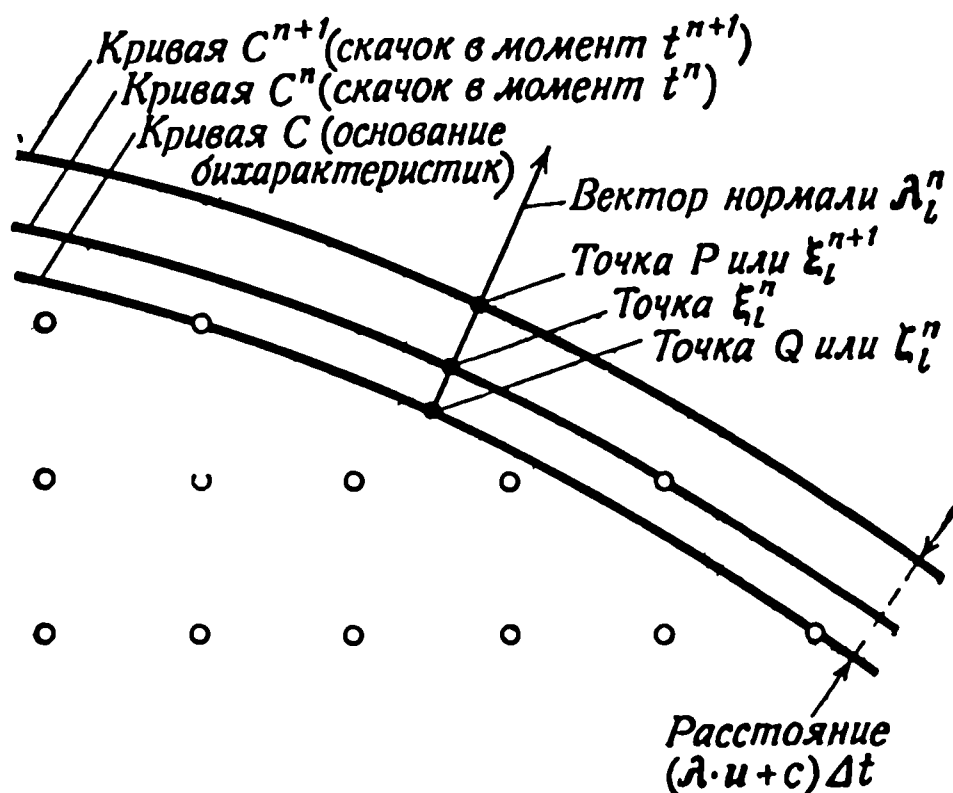
На рис. 13.8 кривые C^n и C^{n+1} изображают два последовательных положения скачка, а кривая C' является геометрическим местом точек, расположенных на расстоянии $(\mathbf{u} \cdot \lambda + c)\Delta t$ позади последнего положения скачка. Если P и Q — точки (ξ_l^{n+1}, t^{n+1}) и (ξ_l^n, t^n) в пространстве координат-времени, где ξ_l^n — пересечение вектора λ_l^n (направленного назад) с кривой C' , то для любой функции f производная $\delta f / \delta t$ может быть аппроксимирована как $[f(P) - f(Q)] / \Delta t$, в то время как $\lambda^* \cdot \nabla$ является оператором дифференцирования по длине дуги вдоль кривой C' в точке Q и обозначается через d/ds . Поэтому уравнение (13.23) аппроксимируется уравнением

$$p(P) - p(Q) + \rho c(Q) \lambda_l^n \cdot (\mathbf{u}(P) - \mathbf{u}(Q)) + ((\Delta / \Delta s) \lambda^* \cdot \mathbf{u})_Q = 0, \quad (13.32)$$

где $\Delta / \Delta s$ — разностное представление оператора d/ds вдоль кривой C' . Для этого уравнения необходимы значения p , ρc и

$\lambda^* \cdot u$ в точках ξ_i^n на кривой C' ; они получаются надлежащей интерполяцией в момент t^n с использованием величин в узлах основной сетки и в точках ξ_i^n скачка.

Чтобы исследовать устойчивость, требуется исключить из алгоритма любой узел сетки, который расположен слишком близко к скачку; поэтому проведем кривую (показанную на рис. 13.7 пунктирно) на фиксированном расстоянии δ позади



Р и с. 13.8. Чертеж, поясняющий уравнение (13.32).

скачка, причем $\delta \sim \Delta x$, и назовем точки сетки, лежащие между этой кривой и скачком, *точками типа A*; значения p_{jk}^n и т. д. в этих точках не используются в расчете; однако когда все другие величины известны в момент $t = (n+1)\Delta t$, интерполированные значения p_{jk}^n и т. д. вычисляются в каждой точке типа A для того, чтобы их использовать позже, если, вследствие движения скачка, точка (j, k) перестанет быть точкой типа A в следующем временном слое.

Методы первого порядка точности, которые обычно более устойчивы, чем второго, могут показаться несовместимыми по точности с методом второго порядка Лакса — Вендроффа, но Стрэнг (частное сообщение) показал, что для задач этого типа сходимость к истинному решению при стремлении шагов сетки к нулю все же остается второго порядка, даже если границы аппроксимируются с первым порядком точности. Это происходит, грубо говоря, потому, что отношение количества граничных точек к количеству всех точек сетки является в пределе малой величиной первого порядка. Такая процедура в любом случае очевидно лучше, чем методы с искусственной вязкостью, в которых границы (скачки) берутся только с нулевым порядком точности.

Любая точка сетки, у которой отсутствует хотя бы одна соседняя точка, называется *точкой типа В* и требует специального исследования. Чтобы рассчитать p_{jk}^{n+1} и другие величины, например, в точке B_5 на рис. 13.7, сначала получаем значения в точке P_1 на скачке путем интерполяции относительно длины дуги, а затем используем P_1 вместо недостающей соседней точки (точки с индексом $j, k + 1$), предварительно незначительно модифицировав формулы Лакса — Вендроффа в связи с неоднородностью пространственной сетки.

Если параметры среды впереди скачка не постоянны, то величины течения непосредственно впереди, обозначенные выше через p_2, ρ_2, u_2, v_2 , будут зависеть от l и n и должны, по-видимому, рассчитываться каким-то образом по точкам сетки впереди. Разработать детали и проверить результаты — это дело будущего.

Попытки осуществить процедуру подгонки для поверхности скольжения, которая может быть контактным разрывом энтропии или поверхностью раздела сред, пока не увенчались успехом. В этом случае имеется девять неизвестных: p, ρ, u, v на каждой из сторон поверхности и нормальная скорость S самой поверхности. Существует три условия на разрыве: p одинаково на обеих сторонах и нормальная составляющая скорости u одна и та же на обеих сторонах и равна S . Поэтому условия на разрыве, по-видимому, должны быть дополнены тремя дифференциальными уравнениями на каждой из сторон. Соответствующие члены поверхностного натяжения должны быть включены в уравнения, чтобы устранить тейлоровскую и гельмгольцеву неустойчивости для волн, длина которых сравнима с Δx .

Многие дополнительные детали двумерных расчетов меняются от задачи к задаче, так что невозможно создать общие или универсальные алгоритмы; специальные алгоритмы должны применяться вблизи точки пересечения трех или более скачков или на пересечении фронта скачка с жесткой стенкой и т. д. Возникает следующий вопрос, носящий в некоторой степени топологический характер: если несколько кривых, представляющих скачки или другие особенности, делят плоскость X, Y в момент t на несколько областей, то необходимо создать подпрограмму для распознавания того, в какой из областей лежит данная точка (j, k) , что можно сделать различными способами. Наконец, для расширяющегося или суживающегося скачка желательно иметь алгоритмы для требующегося иногда увеличения или уменьшения числа точек ξ_j^n на кривой с тем, чтобы сохранять расстояние между ними сравнимым с Δx . Все это сравнительно несложно, тогда как основные процедуры подгонки для внутренних и внешних границ нуждаются в дальнейшем развитии, анализе и проверке.

§ 13.10. Задача о движении атмосферного фронта

Обсуждаемая здесь метеорологическая задача, схематически представленная на рис. 13.9, грубо говоря, является задачей о движении по Земле полярной шапки холодного воздуха. Холодный воздух отделен от теплого воздуха, находящегося выше, довольно отчетливой поверхностью раздела, или *фронтом*, по разные стороны которого температура отличается на $5\text{--}10^\circ\text{C}$. Пересечение фронта с Землей проходит в средних широтах обычно по волнистой линии, называемой на карте погоды «холодным фронтом» или «теплым фронтом» в зависимости от того, что вызывает движущийся фронт: устойчивое похолодание или потепление. Фронт отлого поднимается вверх к северу с углом наклона в (обычно малую) долю градуса, так что на полюсе толщина слоя холодного воздуха достигает нескольких миль.

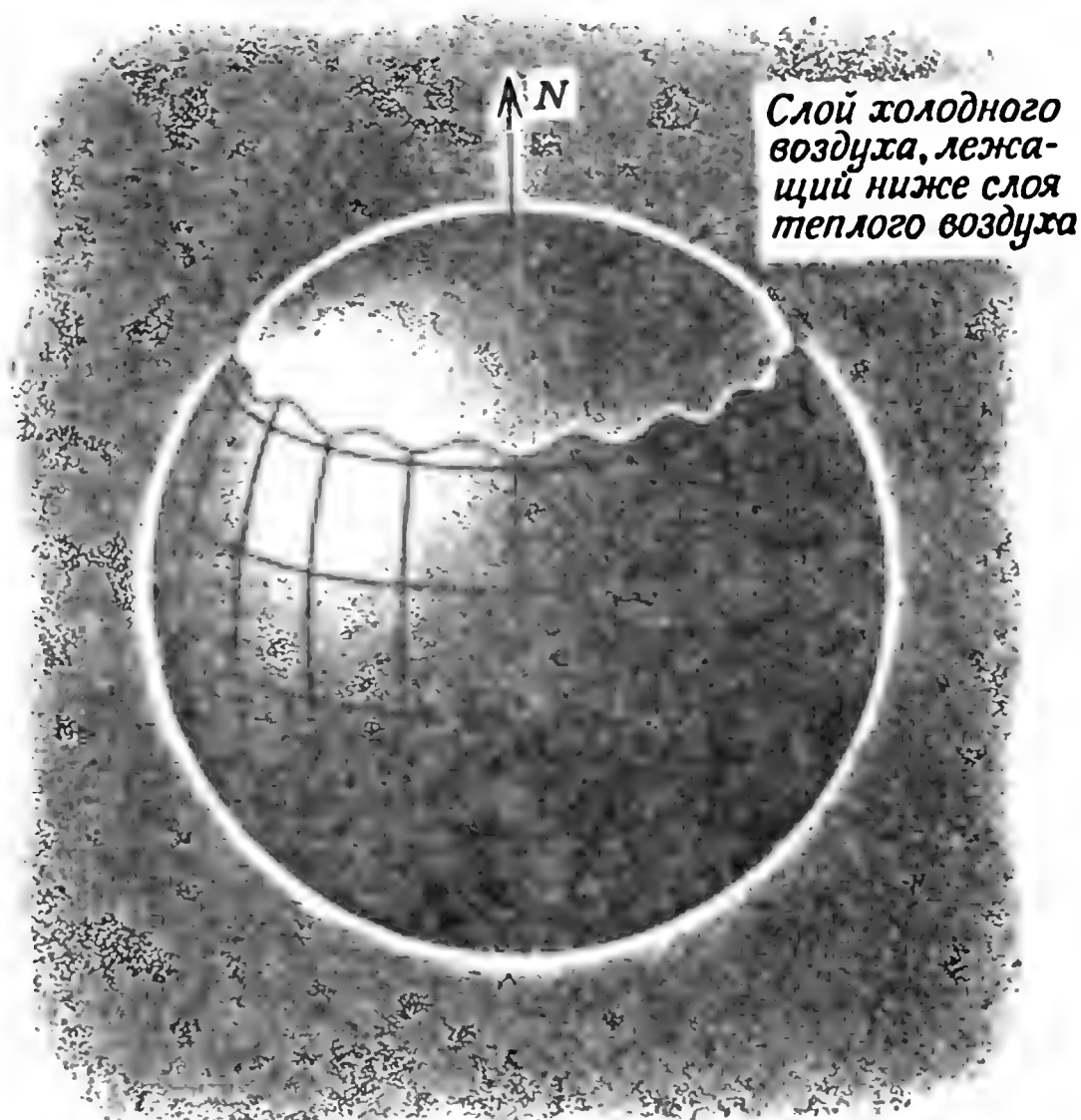


Рис. 13.9. Схематическое изображение холодного фронта.

В простейшем приближении предполагается, что полярная шапка находится в равновесии и кориолисовы силы (обусловленные вращением Земли), уравновешены силой тяжести. В этом приближении считается, что теплый воздух наверху обладает устойчивой зональной циркуляцией с запада на восток относительно холодного воздуха внизу, так что на фронте претерпевает разрыв не только температура, но и широтная компонента скорости ветра. Кориолисова сила, отнесенная к единице объема воздуха, равна $\rho \mathbf{v} \times \boldsymbol{\omega}$, где $\mathbf{v}$ — скорость ветра относительно вращающейся Земли, а $\boldsymbol{\omega}$ — угловая скорость вращения Земли; следовательно, кориолисова сила имеет разные значения для теплых и холодных масс воздуха. С другой стороны, холодный воздух плотнее теплого, так что если бы не было кориолисовых сил, то фронт при гидростатическом равновесии был бы всюду горизонтальным. Равновесие под действием обеих сил обсуждалось Касахарой и др. [1964] и в цитированных

ими работах. В этом приближении, грубо говоря, теплый воздух вращается несколько быстрее вокруг земной оси, чем холодный, и подвергается действию больших центробежных сил; это приводит к тому, что теплый воздух имеет тенденцию собираться на низких широтах и вытеснять холодный воздух в более высокие широты.

Касахара, Изаксон и Стокер [1964, 1965] разработали численный метод расчета движения упрощенной модели фронта как краевой задачи. Полная трехмерная задача с учетом сжимаемости не поддается решению (даже без учета источников тепла, конвекции и тому подобных явлений), но, к счастью, имеются хорошие упрощенные модели. В модели, используемой здесь (Стокер [1953, 1957]), сделаны следующие допущения: 1) расчет производится на достаточно малой части земной поверхности (например, на прямоугольнике размером приблизительно с Соединенные Штаты) с должным образом выбранными дополнительными граничными условиями, так что можно пренебречь кривизной поверхности и изменением кориолисова параметра $f = 2\omega \sin \varphi$, где φ — широта (это приближение иногда называют *приближением касательной плоскости*); 2) вертикальными ускорениями пренебрегают, так что давление в любой точке задается простой гидростатической формулой; 3) сжимаемостью пренебрегают, так как скорость ветра много меньше скорости звука; 4) считается, что каждый слой имеет постоянную плотность; 5) предполагается, что возмущения в холодном слое воздуха пренебрежимо мало влияют на теплый слой, так что его скорость и общая высота двух слоев сохраняют свои равновесные значения в процессе всего расчета. Эти предположения приводят к системе уравнений

$$\frac{Du}{Dt} + g \left(1 - \frac{\rho'}{\rho}\right) \frac{\partial h}{\partial X} = f v, \quad (13.33)$$

$$\frac{Dv}{Dt} + g \left(1 - \frac{\rho'}{\rho}\right) \frac{\partial h}{\partial Y} = f \left(\frac{\rho'}{\rho} \bar{u}' - u\right), \quad (13.34)$$

$$\frac{Dh}{Dt} = 0, \quad (13.35)$$

где D/Dt — оператор

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u \frac{\partial}{\partial X} + v \frac{\partial}{\partial Y},$$

X и Y — декартовы координаты в направлении востока и севера соответственно, $u = u(X, Y, t)$ и $v = v(X, Y, t)$ — компоненты скорости жидкости (ветра) по осям X и Y , $h = h(X, Y, t)$ — вертикальная толщина слоя холодного воздуха, а остальные величины постоянны: ρ'/ρ — отношение плотностей теплого и холодного воздуха, $\bar{u}'$ — постоянная скорость теплого воздуха с запада на восток и g — ускорение силы тяжести.

В смешанной краевой задаче, рассматривавшейся Касахарой и др., течение предполагается периодическим по долготе с таким периодом, что все величины одинаковы в соответственных точках восточной и западной границ прямоугольника, показанного на рис. 13.10 и 13.11; северная граница считается жесткой стенкой; на южной стороне прямоугольника граничных условий не требуется, так как до нее фронт не доходит.

Обозначим через C кривую, по которой фронт пересекает земную поверхность (плоскость X, Y) в момент t , т. е. геометрическое место точек, удовлетворяющих уравнению $h(X, Y, t) = 0$. Если исходить из дифференциальных уравнений, то граничных условий на кривой C не требуется, но дифференциальные уравнения (13.33) — (13.35) должны выполняться на кривой C и выше нее; в численных расчетах удобно размещать точки с координатами $(\xi_l, \eta_l) = \xi_l = \xi_l(t)$ на кривой C , как это показано на рисунках, и следить за движением этих точек и изменением в них величин $(u_l, v_l) = u_l = u_l(t)$, как это делалось при подгонке скачков. Эффективные граничные условия получаются из уравнений (13.33) и (13.34) и имеют вид

$$\frac{d}{dt} \xi_l = v_l, \quad (13.36)$$

$$\frac{d}{dt} v_l = -g \left(1 - \frac{\rho'}{\rho}\right) \nabla h + \gamma v_l + K, \quad (13.37)$$

где

$$\gamma = \begin{bmatrix} 0 & f \\ -f & 0 \end{bmatrix}, \quad K = f \frac{\rho'}{\rho} \bar{u}' \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

и где оператор ∇h должен вычисляться в точке ξ_l .

В методе Касахары и др. используются уравнения Лакса — Вендроффа (с небольшими изменениями) в любой регулярной точке сетки, т. е. в каждой точке (кроме точек северной границы), для которой все восемь соседних точек лежат выше кривой C .

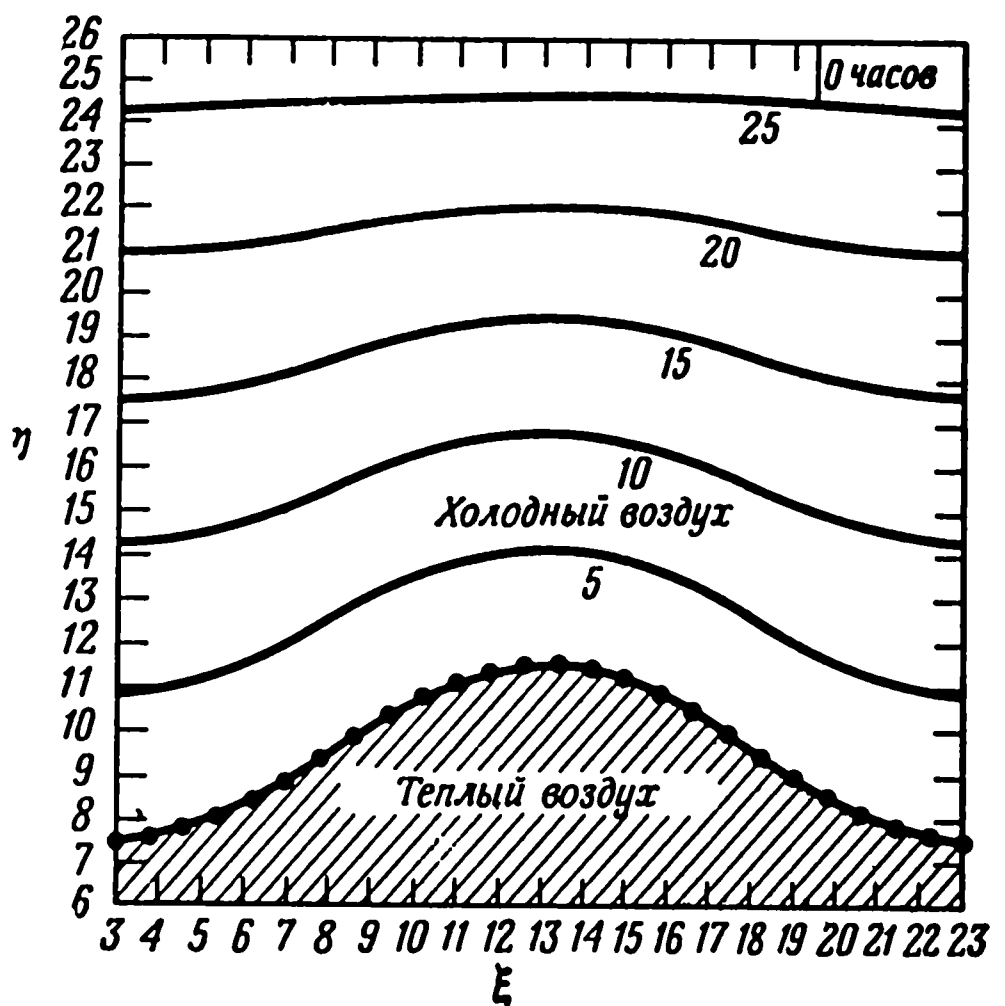


Рис. 13.10. Распределение по вертикали холодного и теплого воздуха в момент $t = 0$.

Точки вблизи кривой C подразделяются на точки типа A и типа B и рассматриваются так же, как при подгонке скачков. Однако Касахара показал возможность и целесообразность проведения всех расчетов со вторым порядком точности. Для этого кривая C рассматривается как последовательность дуг, заданных кубическими уравнениями, которые имеют вид $\eta = f(\xi)$

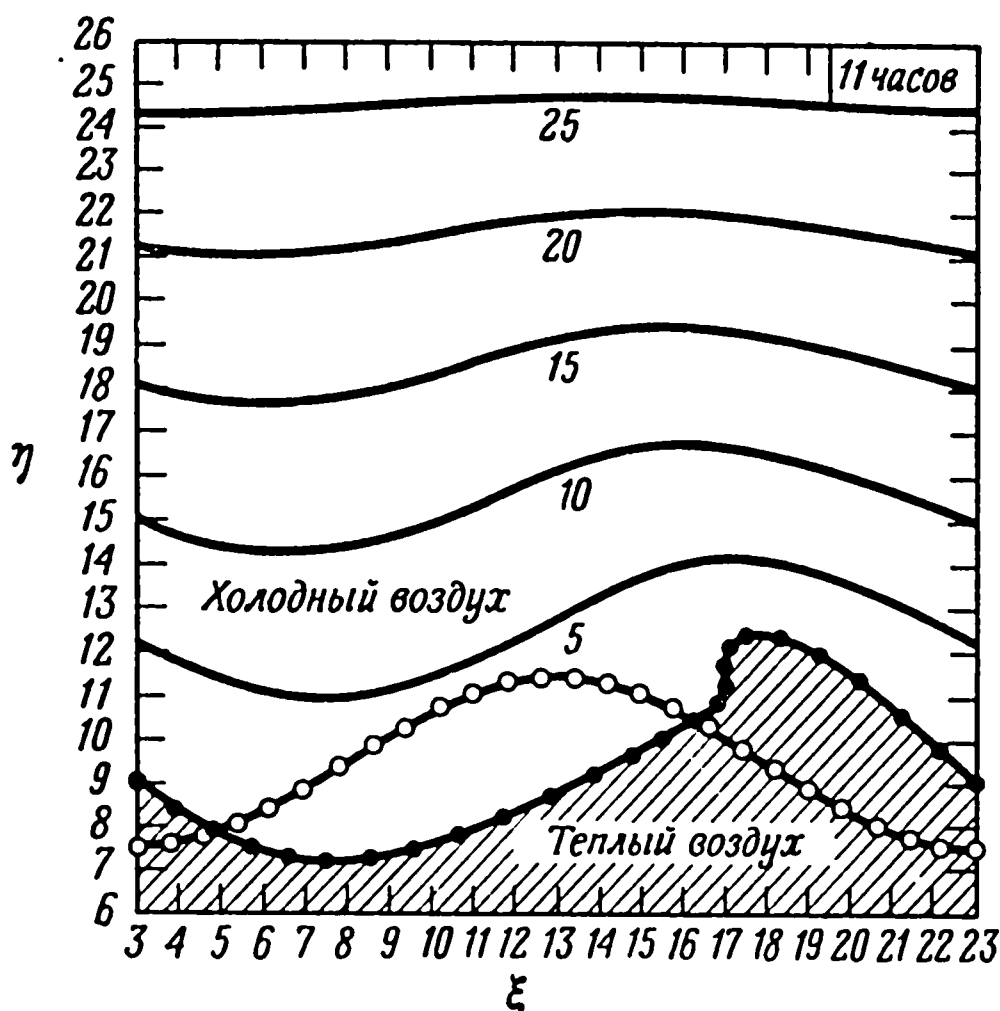


Рис. 13.11. Распределение по вертикали холодного и теплого воздуха в момент $t = 11$ часов.

или $\xi = g(\eta)$ в зависимости от наклона дуги (по-видимому, должно быть достаточно и квадратичного представления), а все интерполяции, и вычисления разностей выполняются с сравнимой точностью. Так как направление вектора ∇h в точке ξ_i на кривой C обязательно совпадает с нормалью λ , то легко провести линию в этом направлении из точки ξ_i до пересечения ее с линией сетки (параллельной оси X или оси Y) в некоторой точке, например Z . С помощью интерполяции на сетке находятся h и $\lambda \cdot \nabla h$ в точке Z ; эти значе-

ния и значение $h = 0$ в точке ξ_i позволяют провести квадратичную подгонку величины h вдоль линии от ξ_i до Z и тем самым определить ∇h в точке ξ_i . Тогда конечноразностная аппроксимация граничных условий (13.36), (13.37) примет вид

$$\xi_i^{n-1} - \xi_i^n = \frac{1}{2}(\mathbf{v}_i^{n+1} + \mathbf{v}_i^n) \Delta t, \quad \mathbf{v}_i^{n+1} - \mathbf{v}_i^n = \frac{1}{2}(\psi_i^{n+1} + \psi_i^n) \Delta t,$$

где ψ — правая часть уравнения (13.37). Как и при подгонке скачков, вся процедура является существенно неявной и для решения уравнений используется итеративный процесс, который подробно описан Касахарой и др. [1964].

В одном из расчетов по этому методу были использованы следующие начальные данные: $h(X, Y, 0)$ бралось, как показано на рис. 13.10, а $u(X, Y, 0) \equiv 0$; рассчитанная конфигурация массы холодного воздуха показана на рис. 13.11. На основании проверки энергии и других величин, сравнения с наблюдениями и т. п. можно сделать заключение, что этот метод устойчив и достаточно точен.

БИБЛИОГРАФИЯ

Звездочкой отмечены работы, добавленные при переводе.
Для переводных книг в квадратных скобках стоит год оригинального издания.

AMS — Amer. Math. Soc.

А д а м а р (H a d a m a r d J.)

[1923] Lectures on Cauchy's problem in linear partial differential equations, Yale Univ. Press.

А н з о р г е (A n s o r g e R.)

[1963] Die Adams-Verfahren als Charakteristikenverfahren höherer Ordnung zur Lösung von hyperbolischen Systemen halbliner Differentialgleichungen, *Numer. Math.*, 5, 443.

А р о н с о н (A r o n s o n D. G.)

[1963] On the stability of certain finite difference approximations to parabolic systems of differential equations, *Numer. Math.*, 5, 118; 290.

Б а г р и н о в с к и й К. А., Г о д у н о в С. К.

[1957] Разностные схемы для многомерных задач, *ДАН*, 115, 431.

Б а н а х (B a n a c h S.)

[1932] Théorie des opérations linéaires, Hafner Publishing Co., N. Y. [Украинский перевод: Курс функционального анализа, ДУПВ «Радянська школа», 1948.]

Б а т л е р (B u t l e r D. S.)

[1960] The numerical solution of hyperbolic systems of partial differential equations in three independent variables, *Proc. Roy. Soc.*, A255, 232.

Б а ч е н а н (B u c h a n a n M. L.)

[1963a] A necessary and sufficient condition for stability of difference schemes for second order initial value problems, *J. Soc. Indust. Appl. Math.*, 11, 474.

[1963b] A necessary and sufficient condition for stability of difference schemes for initial value problems, *J. Soc. Indust. Appl. Math.*, 11, 919.

Б е к к е р (B e c k e r R.)

[1922] Stosswelle und Detonation, *Z. Physik*, 8, 321.

Б е л л м а н (B e l l m a n R.)

*[1960] Введение в теорию матриц, «Наука», 1969.

Белоцерковский О. М.; Чушкин П. И.

- [1962] Численный метод интегральных соотношений, *Ж. выч. мат. и мат. физ.*, 2, № 5, 731.

Бете, Фукс, фон Нейман, Пайерлс, Пенни (Bethe H. A., Fuchs K., von Neumann J., Peierls R., Penny W. G.)

- [1944] Shock hydrodynamics and blast waves, Report AECD-2860, Los Alamos Sci. Lab.

Биркгоф (Birkhoff G.)

- [1962] Helmholtz and Taylor instability, Proceedings of Symposia in Appl. Math., AMS, 13, 55.

Биркгоф, Варга (Birkhoff G., Varga R. S.)

- [1959] Implicit alternating direction methods, *Trans. AMS*, 92, 13.

Блер, Метрополис, фон Нейман, Тауб, Цингоу (Blair A., Metropolis N. C., von Neumann J., Taub A. H., Tsingou M.)

- [1959] A study of a numerical solution to a two-dimensional hydrodynamical problem, *Math. Comp. (MTAC)*, 13, 145.

Бликни, Тауб (Bleakney W., Taub A. H.)

- [1949] Interaction of shock waves, *Rev. Mod. Phys.*, 21, 584.

Бриллюэн (Brillouin M.)

- [1931] Sur quelques problèmes non résolus de la physique mathématique classique. Propagation de la fusion, *Ann. Inst. H. Poincaré*, 1, 285.

Брун, Хаак (Bruhn G., Haack W.)

- [1958] Ein Charakteristikenverfahren für dreidimensionale instationäre Gasströmungen, *ZAMP*, 9b, 173.

Брюс, Писмэн, Ракфорд, Райс (Bruce G. H., Peaceman D. W., Rachford H. H., Jr., Rice J. D.)

- [1953] Calculation of unsteady-state gas flow through porous media, *Trans. Amer. Inst. of Mining and Met. Engrs*, 198, 79.

Бурстейн (Burstein S. Z.)

- [1963] Numerical calculation of multidimensional shocked flows, Report NYO 10433, Courant Inst. Math. Sci., N. Y. Univ.
[1965] Finite difference calculations for hydrodynamic flows containing discontinuities, Report NYO 1480-33, Courant Inst. of Math Sci., N. Y. Univ.

Вазов, Форсайт (Wasow W. R., Forsythe G. E.)

- [1960] Разностные методы для уравнений в частных производных, ИЛ, 1963.

Варга (Varga R. S.)

- [1962] Matrix iterative analysis, Prentice Hall, Englewood Cliffs, N. J.

Вендрофф (Wendroff B.)

- [1960a] On the convergence of the discrete ordinate method, *J. Soc. Indust. Appl. Math.*, 8, 508.
[1960b] On central difference equations for hyperbolic systems, *J. Soc. Indust. Appl. Math.*, 8, 549.

Видлунд (Widlund O. B.)

[1965] On the stability of parabolic difference schemes, *Math. Comp.*, 19, 1.

[1966] Stability of parabolic difference schemes in the maximum norm, *Numer. Math.*, 8, 186.

Вик (Wick G. C.)

[1943] Über ebene Diffusionsprobleme, *Z. Physik*, 121, 702.

Гари (Gary J.)

[1962a] Numerical computation of hydrodynamic flows which contain a shock, Report NYO 9603, Courant Inst. Math. Sci., N. Y. Univ.

[1962b] On certain finite difference schemes for the equations of hydrodynamics, Report NYO 9188, Courant Inst. Math. Sci., N. Y. Univ.

Гельфанд И. М., Шолов Г. Е.

*[1958] Обобщенные функции и действия над ними, Физматгиз.

Гершгорин (Gershgorin S.)

[1931] Über die Abgrenzung der Eigenwerte einer Matrix, *ИАН, сер. физ. и мат.*, 7, 749.

Гёльднер (Göldner S. R.)

[1949] Existence and uniqueness theorems for two systems of nonlinear hyperbolic differential equations for functions of two independent variables, Ph. D. Thesis, N. Y. Univ.

Годунов С. К.

[1959] Разностный метод численного расчета разрывных решений уравнений гидродинамики, *Мат. сб.*, 47, 271.

[1962] Метод ортогональной прогонки для решения систем разностных уравнений, *Ж. выч. мат. и мат. физ.*, 2, № 6, 972.

Годунов С. К., Рябенский В. С.

[1962] Введение в теорию разностных схем, Физматгиз.

[1963a] Спектральные признаки устойчивости краевых задач для несамосопряженных разностных уравнений, *УМН*, 18, 3.

[1963b] Канонические виды систем линейных обыкновенных разностных уравнений с постоянными коэффициентами, *Ж. выч. мат. и мат. физ.*, 3, № 2, 211.

Гординг (Gårding L.)

[1953] Dirichlet problem for linear elliptic partial differential equations, *Math. Scand.*, 1, 53.

Гоуд (Goad W. B.)

[1960] WAT: a numerical method for two-dimensional unsteady fluid flow, Report LAMS 2365, Los Alamos Sci. Lab.

Далквист (Dahlquist G. G.)

[1963] Stability questions for some numerical methods for ordinary differential equations, Proc. Symposia in Appl. Math.; Experimental arithmetic, high speed computing and mathematics, AMS, 147.

Дафф (Duff R. E.)

[1962] Slip line instability, Proceedings of Symposia in Appl. Math., AMS, 13 (Hydrodynamic Instability), 77.

Д ж о н (J o h n F.)

- [1952] On the integration of parabolic equations by difference methods, *Comm. Pure Appl. Math.*, 5, 155.
- [1953] A note on «improper» problems in partial differential equations, *Comm. Pure Appl. Math.*, 8, 591.

Д у г л а с (D o u g l a s J.)

- [1955] On the numerical integration of $\partial^2 u / \partial x^2 + \partial^2 u / \partial y^2 = \partial u / \partial t$ by implicit methods, *J. Soc. Indust. Appl. Math.*, 3, 42.
- [1956] On the relation between stability and convergence in the numerical solution of linear parabolic and hyperbolic equations, *J. Soc. Indust. Appl. Math.*, 4, 20.
- [1957] A note on the alternating direction implicit method for the numerical solution of the heat equation, *Proc. AMS*, 8, 409.
- [1962] Alternating direction methods for three space variables, *Numer. Math.*, 4, 41.

Д у г л а с, Г а н н (D o u g l a s J., G u n n J.)

- [1964] A general formulation of alternating direction methods I, *Numer. Math.*, 6, 428.

Д у г л а с, Р а к ф о р д (D o u g l a s J., R a c h f o r d H. H.)

- [1956] On the numerical solution of and three variables, *Trans. AMS*, 82, 421.

Д ю ф о р т, Ф р а н к е л (d u F o r t E. C., F r a n k e l S. P.)

- [1953] Stability conditions in the numerical treatment of parabolic differential equations, *Math. Tables and other Aids to Computation*, v. 7, 135.

Д ь я к о н о в Е. Г.

- [1962] О некоторых разностных схемах для решения краевых задач, *Ж. вычисл. мат. и мат. физ.*, 2, № 1, 57.
- [1964] Разностные схемы второго порядка точности с расщепляющимся оператором для параболических уравнений без смешанных производных, *Ж. вычисл. мат. и мат. физ.*, 4, № 5, 935.

З а у э р (S a u e r R.)

- [1963] Differenzenverfahren für hyperbolische Anfangswertprobleme bei mehr als zwei unabhängigen Veränderlichen mit Hilfe von Nebencharakteristiken, *Numer. Math.*, 5, 55.

З и г м у н д (Z y g m u n d A.)

- [1952] Тригонометрические ряды, т. 1, 2, «Мир», 1965.

И о г а н с о н, К р а й с (J o h a n n s o n O., K r e i s s H. O.)

- [1963] Über das Verfahren der zentralen Differenzen zur Lösung des Cauchy-Problems für partielle Differentialgleichungen, *Nordisk Tidskr. Informations-Behandling*, 3, 97.

К а л к и н (C a l k i n J. W.)

- [1942] On the spontaneous formation of shocks during a fluid motion, unpublished report.

К а н т о р о в и ч Л. В.

- [1948] Функциональный анализ и прикладная математика, *УМН*, 3, вып. 6, 89.

Карлсон (Carlson B. G.)

- [1953] The S_n method, Unpublished Los Alamos Report and subsequent reports.

Карслоу, Егер (Carslaw H. S., Jaeger J. S.)

- [1964] Теплопроводность твердых тел, «Наука».

Касахара (Kasahara A.)

- [1965] On certain finite-difference methods for fluid dynamics, *Monthly Weather Review*, Jan. 1965, 27.

Касахара, Изаксон, Стокер (Kasahara A., Isaacson E., Stoker J. J.)

- [1964] Numerical studies of frontal motion in the atmosphere. I, Report NYO 1480-6, Courant Inst. Math. Sci., N. Y. Univ.

Като (Kato T.)

- [1960] Estimation of iterated matrices, with application to the von Neumann condition, *Numer. Math.*, 2, 22

Кворлс (Quarles D.)

- [1964] Ph. D. Thesis, N. Y. Univ.

Келлер (Keller H. B.)

- [1956] Difference equations for hyperbolic systems, Seminar, N. Y. Univ.

- [1958] Approximate solutions of transport problems — part I. Steady-state elastic scattering in plane and spherical geometry, *J. Soc. Indust. Appl. Math.*, 6, 452.

- [1960a] Approximate solutions of transport problems — part II. Convergence and applications of the discrete ordinate method, *J. Soc. Indust. Appl. Math.*, 8, 43.

- [1960b] On the pointwise convergence of the discrete ordinate method, *J. Soc. Indust. Appl. Math.*, 8, 560.

- [1960c] Convergence of the discrete ordinate method for the anisotropic scattering transport equation, Symposium on the numerical treatment of ordinary differential equations, integral and integrodifferential equations, Rome; Birkhäuser, Basel, 292.

Келлер, Вендрофф (Keller H. B., Wendroff B.)

- [1957] On the formulation and analysis of numerical methods for time dependent transport equations, *Comm. Pure Appl. Math.*, 10, 567.

Коллатц (Collatz L.)

- [1951] Численные методы решения дифференциальных уравнений, ИЛ, 1953.

Кольски (Kolsky H.)

- [1955] Report LA 1867, Los Alamos Sci. Lab.

Коновалов А. Н.

- [1964] Применение метода расщепления к численному решению динамических задач теории упругости, *Ж. выч. мат. и мат. физ.*, 4, № 4, 760.

Крайс (Kreiss H. O.)

- [1959] Über die Differenzapproximation hoher Genauigkeit bei Anfangswertproblemen für partielle Differentialgleichungen, *Numer. Math.*, 1, 186.

- [1960] Über die Lösung von Anfangswertaufgaben für partielle Differentialgleichungen mit Hilfe von Differenzengleichungen, *Trans. Roy. Inst. Tech.*, Stockholm, 166.
- [1962] Über die Stabilitätsdefinition für Differenzengleichungen die partielle Differentialgleichungen approximieren, *Nordisk Tidskr. Informations-Behandlung*, 2, 153.
- [1963] Über implizite Differenzenmethoden für partielle Differentialgleichungen, *Numer. Math.*, 5, 24.
- [1964] On difference approximations of the dissipative type for hyperbolic differential equations, *Comm. Pure Appl. Math.*, 17, 355.
- [1965] On difference approximations, Symposium on numerical solution of partial differential equations, Univ. Maryland, May, 1965.

К р а н к, Н и к о л с о н (C r a n k J., N i c k o l s o n P.)

- [1947] A practical method for numerical integration of solutions of partial differential equations of heat conduction type, *Proc. Cambridge Philos. Soc.*, 43, 50.

К р е й н М. Г.

- [1958] Интегральное уравнение по полупрямой с ядром, зависящим от разности аргументов, *УМН*, 13, вып. 5, 3.

К р е н д е л л (C r a n d a l l S. H.)

- [1954] Numerical treatment of a fourth-order parabolic differential equation, *J. Assoc. Computing Machinery*, 1, 111.

К у р а н т (C o u r a n t R.)

- [1962] Уравнения с частными производными, «Мир», 1964.

К у р а н т, Г и л ь б е р т (C o u r a n t R., H i l b e r t D.)

- [1953] Методы математической физики, т. 1, ГИТТЛ, 1951 (перевод с предыдущего издания).

К у р а н т, И з а к с о н, Р и с (C o u r a n t R., I s a a c s o n E., R e e s M.)

- [1952] On the solutions of nonlinear hyperbolic differential equations by finite differences, *Comm. Pure Appl. Math.*, 5, 243.

К у р а н т, Л а к с (C o u r a n t R., L a x P. D.)

- [1949] On nonlinear partial differential equations with two independent variables, *Comm. Pure Appl. Math.*, 2, 255.

К у р а н т, Ф р и д р и х с (C o u r a n t R., F r i e d r i c h s K. O.)

- [1948] Сверхзвуковые течения и ударные волны, ИЛ, 1950.

К у р а н т, Ф р и д р и х с, Л е в и (C o u r a n t R., F r i e d r i c h s K. O., L e w y H.)

- [1928] Über die partiellen Differenzengleichungen der mathematischen Physik, *Math. Ann.*, 100, 32. [Русский перевод: О разностных уравнениях математической физики, *УМН*, вып. VIII (1940), 125.]

Л а а с о н е н (L a a s o n e n P.)

- [1949] Über eine Methode zur Lösung der Wärmeleitungsgleichung, *Acta Math.*, 81, 309.

Л а д ф о р д, П о л я ч е к, З е г е р (L u d f o r d G. S. S., P o l a c h e k R. J., S e e g e r R.)

- [1953] On unsteady flow of viscous fluids, *J. Appl. Phys.*, 24, 490.

Ладыженская О. А.

- [1952] Решение задачи Коши для гиперболических систем методом конечных разностей, *Уч. зап. ЛГУ*, сер. мат., 144, № 23, 192.
- [1953] Смешанная задача для гиперболического уравнения, ГИТТЛ.
- [1957a] Метод конечных разностей в теории уравнений с частными производными, *УМН*, вып. 5, 123.
- [1957b] О построении разрывных решений квазилинейных гиперболических уравнений как пределов решений соответствующих параболических уравнений при стремлении «коэффициента вязкости» к нулю, *Труды Моск. мат. о-ва*, 6, 465.

Лакс А. (Лакс А.)

- [1956] On Cauchy's problem for partial differential equations with multiple characteristics, *Comm. Pure Appl. Math.*, 6, 135.

Лакс П. (Лакс Р.)

- [1953] Nonlinear hyperbolic equations, *Comm. Pure Appl. Math.*, 6, 231.
- [1954] Weak solutions of nonlinear hyperbolic equations and their numerical computation, *Comm. Pure Appl. Math.*, 7, 159.
- [1957] Hyperbolic system of conservation laws. II, *Comm. Pure Appl. Math.*, 10, 537.
- [1960] The scope of the energy method, *Bull. AMS*, 66, 32.
- [1961] On the stability of difference approximations to solutions of hyperbolic equations with variable coefficients, *Comm. Pure Appl. Math.*, 14, 497.

Лакс, Вендрофф (Лакс Р. Д., Вендрофф В.)

- [1960] Systems of conservation laws, *Comm. Pure Appl. Math.*, 13, 217.
- [1962] On the stability of difference schemes, *Comm. Pure Appl. Math.*, 15, 363.
- [1964] Difference schemes for hyperbolic equations with high order of accuracy, *Comm. Pure Appl. Math.*, 17, 381.

Лакс, Келлер (Лакс Р. Д., Келлер Дж. В.)

- [1951] The initial and mixed boundary value problem for hyperbolic systems, Report LAMS-1205, Los Alamos Sci. Lab.

Лакс, Ниренберг (Лакс Р. Д., Ниренберг Л.)

- [1966] On stability for difference schemes: a sharp form of Gårding's inequality, *Comm. Pure Appl. Math.*, 19, № 4, 437—492.

Лакс, Рихтмайер (Лакс Р. Д., Рихтмайер Р. Д.)

- [1956] Survey of the stability of linear finite difference equations, *Comm. Pure Appl. Math.*, 9, 267.

Левин (Левин Н.)

- [1927] Über das Anfangswertproblem bei einer hyperbolischen nichtlinearen partiellen Differentialgleichung zweiter Ordnung mit zwei unabhängigen Veränderlichen, *Math. Ann.*, 98, 179.
- [1948] On the convergence of solutions of difference equations, Essays presented to R. Courant on his 60th birthday, Interscience Publishers, N. Y., 211.

Лейт (Лейт С. Е.)

- [1964] Lagrangian advection in an atmospheric model, Report UCRL 7822, Lawrence Radiation Lab., Univ. California, Livermore.

Ленер, Уинг (Lehner J., Wing G. M.)

- [1955] On the spectrum of an unsymmetric operator arising in the transport theory of neutrons, *Comm. Pure Appl. Math.*, 8, 217.

Либман (Liebmann L.)

- [1918] Die Angenäherte Ermittlung harmonischer Funktionen und konformer Abbildungen, *Sitz. Bayer Akad. Wiss., Math.-Phys. Klasse*, 385.

Лиз (Lees M.)

- [1960a] A priori estimates for the solution of difference approximations to parabolic partial differential equations, *Duke Math. J.*, 27, 297.
 [1960b] Energy inequalities for the solution of differential equations, *Trans. AMS*, 94, 58.
 [1961] Alternating direction and semi-explicit difference methods for parabolic partial differential equations, *Numer. Math.*, 3, 398.

Лоджмэн (Logemann G.)

- [1965] Existence and uniqueness of rarefaction waves, Ph. D. thesis, N. Y. Univ.

Льютерт (Leutert W.)

- [1951] On the convergence of approximate solutions of the heat equation to the exact solution, *Proc. AMS*, 2, 433.
 [1952] On the convergence of unstable approximate solutions at the heat equation to the exact solution, *J. Math. Physics*, 30, 245.

Мак-Даффн (MacDuffee C. C.)

- [1946] The theory of matrices, Chelsea Publishing, Bronx, N. Y.

Мартин (Martin M. H.)

- [1955] An existence and uniqueness theorem for unsteady one-dimensional anisentropic flow, *Comm. Pure Appl. Math.*, 8, 367.

Марчук Г. И., Яненко Н. Н.

- [1966] Применение метода расщепления (дробных шагов) для решения задач математической физики, сб. «Некоторые вопросы вычислительной и прикладной математики», Новосибирск, 5.

Миллер, Стрэнг (Miller J. J. H., Strang W. G.)

- [1965] Matrix theorems for partial differential and difference equations, Stanford University Tech. Report CS28, Stanford, California.

Милн (Milne W. E.)

- [1953] Численные методы решения дифференциальных уравнений, ИЛ, 1955.

Митчелл, Пирс (Mitchell A. R., Pearce R. P.)

- [1962] High accuracy difference formulae for the numerical solution of the heat conduction equation, *Computer J.*, 5, 142.

Моравец (Morawetz C. S.)

- [1956] On the non-existence of continuons transonic flows past profiles, *Comm. Pure Appl. Math.*, 9, 45.

Моримото (Morimoto H.)

- [1962] Stability in the wave equation coupled with heat flow, *Numer. Math.*, 4, 136.

Мортон (Morton K. W.)

[1964a] On a matrix theorem due to H. O. Kreiss, *Comm. Pure Appl. Math.*, 17, 375.

[1964b] Finite amplitude compression waves in a collision free plasma. *Physics of Fluids*, 7, 1800.

Мортон, Шехтер (Morton K. W., Schechter S.)

[1965] On the stability of finite difference matrices, *J. Soc. Indust. Appl. Math. Numer. Anal.*, ser. B, 2, 119.

Нейман (von Neumann J.)

[1943] Математические основы квантовой механики, «Наука», 1964.

Нейман, Рихтмайер (von Neumann J., Richtmyer R. D.)

[1950] A method for the numerical calculations of hydrodynamical shocks, *J. Appl. Phys.*, 21, 232. [Русский реферат: Метод численного расчета гидродинамических скачков, сб. «Механика», № 1 (1951), 27.]

Нортон (Van Norton R.)

[1962] On the real spectrum of a mono-energetic neutron transport operator, *Comm. Pure Appl. Math.*, 15, 99.

Нох (Noh E. F.)

[1964] CEL: A time-dependent two-space-dimensional, coupled Eulerian-Lagrangian code, *Methods in Comput. Phys.*, Academic Press, N. Y., 3, 117. [Русский перевод: СЭЛ — совместный эйлерово-лагранжев метод для расчета нестационарных двумерных задач, Сб. «Вычислительные методы в гидродинамике», «Мир», 1967, 128.]

О'Брайен, Хаймэн, Каплан (O'Brien G. G., Human M. A., Kaplan S.)

[1951] A study of the numerical solution of partial differential equations, *J. Math. Physics*, 29, 223.

Олейник О. А.

[1956] Задачи Коши для нелинейных дифференциальных уравнений первого порядка с разрывными начальными условиями, Труды Моск. матем. о-ва, 5, 433.

[1957] Разрывные решения нелинейных дифференциальных уравнений, УМН, 12, вып. 3, 3.

Олейник О. А., Введенская Н. Д.

[1957] Решение задачи Коши и краевой задачи для нелинейных уравнений в классе разрывных функций, ДАН, 113, 503.

Парлетт (Parlett B. N.)

[1966] Accuracy and dissipation in difference schemes, *Comm. Pure. Appl. Math.*, 19, 111.

Партер (Parter S. V.)

[1962] Stability, convergence, and pseudo-stability of finite difference equations for an over-determined problem, *Numer. Math.*, 4, 277.

Пирс, Митчелл (Pearce R. P., Mitchell A. R.)

[1962] On finite difference methods of solution of the transport equations, *Math. Comp.*, 16, 155.

Пирси (Pearcy G.)

- [1966] An elementary proof of the power inequality for the numerical radius, *Michigan Math. J.*, 13, № 3, 289—291.

Писмэн, Ракфорд (Peaceman D. W., Rachford H. H. Jr.)

- [1955] The numerical solution of parabolic and elliptic differential equations, *J. Soc. Industrial Appl. Math.*, 3, 28.

Рисс, Секефальви-Надь

- [1952] Лекции по функциональному анализу, ИЛ, 1954.

Рихтмайер (Richtmyer R. D.)

- [1960] Proposed fluid-dynamics research, Unpublished memorandum, Courant Inst. Math. Sci., N. Y. Univ.
[1961] Progress report on the Mach reflection calculation, Report NYO 9764, Courant Inst. Math. Sci., N. Y. Univ.
[1962] A survey of difference methods for non-steady fluid dynamics, NCAR Technical note 63-2, Nat'l Center for Atmos. Res., Boulder, Colo.
[1964] The stability criterion of Godunov and Ryabenkii for difference schemes, Report NYO 1840-4, Courant Inst. Math. Sci., N. Y. Univ.

Рихтмайер, Мортон (Richtmyer R. D., Morton K. W.)

- [1964] Stability studies for difference equations: (I) Non-linear instability, (II) Coupled sound and heat flow, Report NYO 1480-5, Courant Inst. Math. Sci., N. Y. Univ.

Ричардсон (Richardson D. J.)

- [1964] The solution of two-dimensional hydrodynamic equations by the method of characteristics, *Methods in Comput. Phys.*, Academic Press, N. Y., v. 3, 295. [Русский перевод: Метод характеристик для решения уравнений гидродинамики двумерных неустановившихся течений, сб. «Вычислительные методы в гидродинамике», «Мир», 1967, 292.]

Робертс, Вайс (Roberts K. V., Weiss N. O.)

- [1966] Convective difference schemes, *Math. Comp.*, 20, 272.

Рождественский Б. Л.

- [1960] Разрывные решения систем квазилинейных уравнений гиперболического типа, *УМН*, 15, вып. 6, 59.

Роуз (Rose M. E.)

- [1956] On the integration of non-linear parabolic equations by implicit difference methods, *Quart. Appl. Math.*, 14, 237.

Рябенский В. С.

- [1964] Необходимые и достаточные условия хорошей обусловленности краевых задач для систем обыкновенных разностных уравнений, *Ж. вычисл. мат. и мат. физ.*, 4, № 2, 242.

Рябенский В. С., Филиппов А. Ф.

- [1956] Об устойчивости разностных уравнений, ГИТТЛ.

Самарский А. А.

- [1964] Экономичные разностные системы для систем уравнений параболического типа, *Ж. вычисл. мат. и мат. физ.*, 4, № 5, 227.

Санто, Келлер (de Santo D. F., Keller H. B.)

- [1961] Numerical studies of transition from laminar to turbulent flow over a flat plate, Coll. Eng. Res. Div. Report, N. Y. Univ.

Саульев В. К.

- [1957] Об одном способе численного интегрирования уравнений диффузии, *ДАН*, 115, 1077.
[1958] О методах повышенной точности и двухсторонних приближениях к решению параболических уравнений, *ДАН*, 118, 1088.

Саусвелл (Southwell R. V.)

- [1946] Relaxation methods in theoretical physics, Oxford University Press.

Сейдман (Seidman T. J.)

- [1963] On the stability of certain difference schemes, *Numer. Math.*, 5, 201.

Софронов И. Д.

- [1963] К разностному решению уравнения теплопроводности в криволинейных координатах, *Ж. вычисл. мат. и мат. физ.*, 3, № 1, 786.

Спицер (Spitzer L., Jr.)

- [1962] Физика полностью ионизованного газа, «Мир», 1965.

Старк (Stark R. H.)

- [1956] Rates of convergence in numerical solution of the diffusion equation, *J. Assoc. Comput. Mach.*, 3, 29.

Стеттер (Stetter H. J.)

- [1961] On the convergence of characteristic finite-difference methods of high accuracy for quasi-linear hyperbolic equations, *Numer. Math.*, 3, 321.

Стокер (Stoker J. J.)

- [1953] Dynamical theory for treating the motion of cold and warm fronts in the atmosphere, Report IMM 200, Courant Inst. Math. Sci., N. Y. Univ.
[1957] Волны на воде, ИЛ, 1959.

Стоун (Stone M. H.)

- [1932] Linear transformations in Hilbert space, AMS Colloq. Publications.

Стрэнг (Strang W. G.)

- [1960] Difference methods for mixed boundary-value problems, *Duke Math. J.*, 27, 221.
[1962] Polynomial approximation of Bernstein type, *Trans. AMS*, 105, 525.
[1963] Accurate partial difference methods. I. Linear Cauchy problems, *Arch. Rational Mech. Anal.*, 12, 392.
[1964a] Accurate partial difference methods. II. Nonlinear problems, *Numer. Math.*, 6, 37.
[1964b] Wiener — Hopf difference equations, *J. Math. Mech.*, 13, 85.
[1965] Necessary and insufficient conditions for well posed Cauchy problems, *J. Differential Equations*, 2, 107.

Тёрнер, Вендрофф (Turner J., Wendroff B.)

- [1964] An unconditionally stable implicit difference scheme for the hydrodynamical equations, Report LA-3007, Los Alamos Sci. Lab.

Тёрниг (Törnig W.)

- [1963] Über Differenzenverfahren in Rechteckgittern zur numerischen Lösung quasi-linearer hyperbolischer Differentialgleichungen, *Numer. Math.*, 5, 353.

Тодд (Todd J.)

- [1955] A direct approach to the problem of stability in the numerical solution of partial differential equations, Report № 4260, National Bureau of Standards.

Томас (Thomas L. H.)

- [1944] Note on Becker's theory of the shock front, *J. Chem. Phys.*, 12, 449.
[1949] Numerical solution of partial differential equations of parabolic type, Seminar of Scientific Computation, International Business Machines Corp.
[1950] Stability of partial difference equations, Symposium on theoretical compressible flow, Report № 1132, U. S. Naval Ordnance Laboratory.
[1954] Computation of one-dimensional compressible flows including shocks, *Comm. Pure Appl. Math.*, 7, 195.

Томпсон Р. (Thompson R. J.)

- [1964] Difference approximation for inhomogeneous and quasi-linear equations, *J. Soc. Indust. Appl. Math.*, 12, 189.

Томпсон У. (Thompson W. B.)

- [1962] An introduction to plasma physics, Pergamon Press, Oxford.

Томи (Thomé V.)

- [1962] A stable difference scheme for the mixed boundary problem for a hyperbolic first order system in two dimensions, *J. Soc. Indust. Appl. Math.*, 10, 229.
[1965] Stability of difference schemes in the maximum norm, *J. Differential Equations*, 1, 273.
[1966] Parabolic difference operators, *Math. Scand.*, 19, № 1, 77—107.

Трулио (Trulio J. G.)

- [1964] The strip code and the jetting of gas between plates, Method in Comput. Phys., Academic Press, N. Y., v. 3, 69. [Русский перевод: Метод полос и течение газа между пластинами, сб. «Вычислительные методы в гидродинамике», «Мир», 1967, 76.]

Тэлбот (Talbot G. P.)

- [1963] Application of the numerical method of characteristics in three independent variables to shockthermal layer interaction problems, Report, Royal Armament Research and Development, U. K.

Уиттекер, Ватсон (Whittaker E. T., Watson G. N.)

- [1927] Курс современного анализа, т. 1, 2, Физматгиз, 1963.

Урм В. Я.

- [1961] О необходимых и достаточных условиях устойчивости разностных уравнений, *ДАН*, 139, 40.

Филлипс Н. (Phillips N. A.)

- [1959] An example of non-linear computational instability, The Atmosphere and the sea in Motion, Bert Bolin, ed., Rockefeller Institute Press, N. Y. 501.

Филлипс Р. (Phillips R. S.)

[1957] Dissipative hyperbolic systems, *Trans. AMS*, 86, 109.

[1959] Dissipative operators and hyperbolic systems of partial differential equations, *Trans. AMS*, 90, 193.

Фландерс, Шортли (Flanders D. A., Shortley G.)

[1950] Numerical determination of fundamental modes, *J. Appl. Phys.*, 21, 1326.

Франк, Лазарус (Frank R. M., Lazarus R. B.)

[1964] Mixed Eulerian Lagrangian method, *Methods in Comput. Phys.*, Academic Press, N. Y., v. 3, 47. [Русский перевод: Смешанный метод, использующий переменные Эйлера и Лагранжа, сб. «Вычислительные методы в гидродинамике», «Мир», 1967, 55.]

Франкел (Frankel S. P.)

[1950] Convergence rates of iterative treatments of partial differential equations, *Math. Tables and other Aids to Computation*, v. 4, 65.

Фридрихс (Friedrichs K. O.)

[1954] Symmetric hyperbolic linear differential equations, *Comm. Pure Appl. Math.*, 7, 345.

[1958] Symmetric positive linear differential equations, *Comm. Pure Appl. Math.*, 11, 333.

Фромм (Fromm J. E.)

[1964] The time-dependent flow of an incompressible fluid, *Methods in Comput. Phys.*, Academic Press, N. Y., v. 3, 346. [Русский перевод: Неустановившееся течение несжимаемой вязкой жидкости, сб. «Вычислительные методы в гидродинамике», «Мир», 1967, 343.]

Хайн К., Хайн Г., Робертс К., Робертс С., Кёппендёрфер (Hain K., Hain G., Roberts K. V., Roberts S. J., Köppendörfer W.)

[1960] Fully ionized pinch collapse, *Z. Naturforsch.*, 15a, 1039.

Харлоу (Harlow F. H.)

[1964] The particle-in-cell computing method for fluid dynamics, *Methods in Comput. Phys.*, Academic Press, N. Y., v. 3, 319. [Русский перевод: Численный метод частиц в ячейках для задач гидродинамики, сб. «Вычислительные методы в гидродинамике», «Мир», 1967, 316.]

Херш (Hersch R.)

[1963] Mixed problems in several variables, *J. Math. Mech.*, 12, 317.

Хилдебранд (Hildebrand F. B.)

[1952] On the convergence of numerical solutions of the heat-flow equations, *J. Math. Physics*, 31, 35.

Хопф (Hopf E.)

[1950] The partial differential equation $u_t + uu_x = \mu u_x$, *Comm. Pure Appl. Math.*, 3, 201.

Хоскин (Hoskin N. E.)

- [1964] Solution by characteristics of the equations of one-dimensional unsteady flow, *Methods in Comput. Phys.*, Academic Press, N. Y., v. 3, 365. [Русский перевод: Метод характеристик для решения уравнений гидродинамики двумерных неустановившихся течений, сб. «Вычислительные методы в гидродинамике», «Мир», 1967, 264.]

Хункоса, Янг (Juncosa M. L., Young D. M.)

- [1953] On the order of convergence of solutions of a difference equation to a solution of the diffusion equation, *J. Soc. Indust. Appl. Math.*, 1, 111.

Чандрасекхар (Chandrasekhar S.)

- [1944a] On the radiative equilibrium of a stellar atmosphere. I, *Astroph. J.*, 99, 180.
[1944b] On the radiative equilibrium of a stellar atmosphere. II, *Astroph. J.*, 100, 82.

Шолов Г. Е.

- *[1961] Математический анализ, Специальный курс, Физматгиз.
*[1965] Математический анализ, Второй специальный курс, «Наука».

Шортли (Shortley G.)

- [1952] Use of Tschebyscheff-polynomial operators in the numerical solution of boundary-value problems, *J. Appl. Phys.*, 24, 392.

Шортли, Уэллер (Shortley G., Weller R.)

- [1938] The numerical solution of Laplace's equation, *J. Appl. Phys.*, 9, 334.

Шultz (Schultz W. D.)

- [1964a] Tensor artificial viscosity for numerical hydrodynamics, *J. Math. Phys.*, 5, 133.
[1946b] Two-dimensional Lagrangian hydrodynamical difference equations, *Methods in Comput. Phys.*, Academic Press, N. Y., v. 3, 1. [Русский перевод: Двумерные конечноразностные гидродинамические уравнения в переменных Лагранжа, сб. «Вычислительные методы в гидродинамике», «Мир», 1967, 9.]

Шур (Schur I.)

- [1909] Über die charakteristischen Wurzeln einer linearen Substitution mit einer Anwendung auf Theorie der Integralgleichung, *Math. Ann.*, 66, 488.

Эванс, Бруссо, Кирстед (Evans G. W., Brousseau R., Kierstead R.)

- [1954] Instability considerations for various difference equations derived from the diffusion equation, Stanford Research Institute, Stanford, California.

Эдди (Eddy R. P.)

- [1949] Stability in the numerical solution of initial value problems in partial differential equations, Naval Ordnance Laboratory Memo. 10232.

Эллиот (Elliott L. A.)

- [1962] Shock fronts in two-dimensional flow, *Proc. Roy. Soc.*, A 267, 558.

Эрлих, Гурвиц (Ehrlich R., Hurwitz H. Jr.)

[1954] Multigroup methods for neutron problems, *Nucleonics*, 12, 23.

Эшенхерст, Метрополис (Ashenurst R. L., Metropolis N. C.)

[1959] Unnormalized floating-point arithmetic, *J. Assoc. Comput. Mach.*, 6, 415.

Янг (Young D. M.)

[1954] Iterative methods for solving partial differential equations of elliptic type, Doctoral Disseration, Harvard University (1950), *Trans. AMS*, 76, 92.

Яненко Н. Н.

[1959] Об одном разностном методе счета многомерного уравнения теплопроводности, *ДАН*, 125, 1207.

[1964] О слабой аппроксимации систем дифференциальных уравнений, *Сиб. матем. ж.*, 5, № 6, 1431.

*[1967] Метод дробных шагов решения многомерных задач математической физики, «Наука», Сибирск. отделение, Новосибирск.

ДОПОЛНИТЕЛЬНАЯ БИБЛИОГРАФИЯ

А б р а м о в А. А.

- [1961a] Вариант метода прогонки, *Ж. вычисл. матем. и матем. физ.*, 1 : 2, 349—351.
- [1961б] О переносе граничных условий для систем линейных обыкновенных дифференциальных уравнений (вариант метода прогонки), *Ж. вычислит. матем. и матем. физ.*, 1 : 3, 542—545.

А б р а м о в А. А., Г а и п о в а А. Н.

- [1971] О численном решении некоторых систем для задач типа Стефана, *Ж. вычислит. матем. и матем. физ.*, 11 : 1, 122—128.

А д а м с к а я И. А.

- [1963] Запись кинетического уравнения в сферических гармониках (криволинейные координаты — осесимметричный случай), *Ж. вычислит. матем. и матем. физ.*, 3 : 5, 927—941.

А д а м с к а я И. А., Г о д у н о в С. К., П р о к о п о в Г. П.

- [1968] Выражение функционала В. С. Владимирова через сферические гармоники в тензорной форме (P_3 -приближение), *Ж. вычислит. матем. и матем. физ.*, 2 : 4, 824—835.

А л а л ы к и н Г. Б., Г о д у н о в С. К., К и р е е в а И. Л., П л и н е р Л. А.

- [1970] Решение одномерных задач газовой динамики в подвижных сетках, М., «Наука».

А н д р е е в В. Б.

- [1966] О равномерной сходимости некоторых разностных схем, *Ж. вычислит. матем. и матем. физ.*, 6 : 2, 238—250.
- [1967a] О разностных схемах с расщепляющимся оператором для общих p -мерных параболических уравнений второго порядка со смешанными производными, *Ж. вычислит. матем. и матем. физ.*, 7 : 2, 312—321.
- [1967б] Об одном методе численного решения третьей краевой задачи для уравнения параболического типа в p -мерном параллелепипеде, *Сб. работ ВЦ МГУ*, 6, 64—75.
- [1969] О сходимости разностных схем с расщепляющимся оператором, аппроксимирующих третью краевую задачу для параболического уравнения, *Ж. вычислит. матем. и матем. физ.*, 9 : 2, 337—349.

А н у ч и н а Н. Н.

- [1970] О методах расчета течений сжимаемой жидкости с большими деформациями, информ. бюллет. «Численные методы механики сплошной среды», 1 : 4, Новосибирск, 3—10.

Бабенко К. И., Молчанов А. М., Русанов В. В., Шноль Э. Э.

- [1963] Методы решения некоторых двумерных задач, сб. «Вопросы вычислит. матем. и вычислит. техники», М., 99—103.

Баклановская В. Ф.

- [1961] Численное решение одномерной задачи для уравнений типа нестационарной фильтрации, *Ж. вычислит. матем. и матем. физ.*, 1 : 3, 461—469.
[1964] Исследование метода сеток решения первой краевой задачи для уравнения типа нестационарной фильтрации, сб. «Численные методы решений дифференциальных и интегральных уравнений», М., 228—243.

Баклановская В. Ф., Гаипова А. Н.

- [1966] Об одной двумерной задаче нелинейной фильтрации, сб. «Численные методы решения задач математической физики», М., 237—241.

Баклановская В. Ф., Кутасов А. Д., Нейгауз М. Г.

- [1961] Численное нахождение квазистационарного решения системы одномерных параболических уравнений, *Ж. вычислит. матем. и матем. физ.*, 1 : 2, 354—357.

Баталова М. В., Бахрах С. М., Винокуров С. А., Загусткин В. Л., Иванов Л. Н., Калманович А. И., Шиндерман И. Д.

- [1969] Комплекс «Сигма» для расчета задач двумерной газодинамики, Труды всесоюзного семинара по численным методам механики вязкой жидкости, Новосибирск, «Наука», 283—292.

Бахвалов Н. С.

- [1962а] О накоплении вычислительной погрешности при численном решении дифференциальных уравнений, *Сб. работ ВЦ МГУ*, 1, 47—68.
[1962б] Об оптимальных способах задания информации при решении дифференциальных уравнений, *Ж. вычислит. матем. и матем. физ.*, 2 : 4, 569—592.
[1966] Об асимптотике при малых ε решения уравнения $u_t + (\varphi(u))_x = \varepsilon u_{xx}$, соответствующего волне разрежения, *Ж. вычислит. матем. и матем. физ.*, 6 : 3, 521—526.
[1967] О параболических системах с малыми параметрами при старших производных, *ДАН СССР*, 174 : 2, 263—266.
[1969] К оптимизации методов решения краевых задач при наличии пограничного слоя, *Ж. вычислит. матем. и матем. физ.*, 9 : 4, 841—859.
[1970] О свойствах оптимальных методов решения задач математической физики, *Ж. вычислит. матем. и матем. физ.*, 10 : 3, 555—568.

Белоцерковский О. М.

- [1969] Численные методы решения задач механики сплошных сред, М., изд. ВЦ АН СССР.
[1970] Обтекание затупленных тел сверхзвуковым потоком газа, Теорет. экспер. иссл., изд. 2-е, М., изд. ВЦ АН СССР.

Белоцерковский О. М., Давыдов Ю. М.

- [1970] Нестационарный метод «крупных частиц» для решения задач внешней аэродинамики, М., изд. ВЦ АН СССР.

Белоцерковский О. М., Чушкин П. И.

- [1962] Численный метод интегральных соотношений, *Ж. вычислит. матем. и матем. физ.*, 2 : 5, 731—759.

Белухина И. Г.

- [1965] О разностной аппроксимации некоторых нелинейных граничных условий для одномерных параболических уравнений, *Сб. работ ВЦ МГУ*, 3, 223—231.
[1969] Разностные схемы для решения плоской динамической задачи теории упругости со смешанными краевыми условиями, *Ж. вычислит. матем. и матем. физ.*, 9 : 3, 362—372.

Березин И. С., Жидков Н. П.

- [1962] Методы вычислений, т. 2, изд. 2-е, М., Физматгиз.
[1963] Методы вычислений, т. 1, изд. 3-е, М., «Наука».

Бондаренко П. С.

- [1964] Вычислительные алгоритмы приближенного решения операторных уравнений, *ДАН УССР*, 3, 295—299.
[1966] Принципы построения оптимальных алгоритмов, *Вычислительная математика*, Киев, 2, 66—77.

Бояринцев Ю. Е.

- [1965] О сходимости разностных схем для волнового уравнения с переменными коэффициентами, *ДАН СССР*, 165 : 3, 474—475.
[1966a] О сходимости метода расщепления и о локальном критерии корректности для разностных уравнений с переменными коэффициентами, *сб. «Некоторые вопросы вычислит. и прикладной матем.»*, Новосибирск, 92—100.
[1966b] О сходимости разностных схем для уравнений с переменными коэффициентами, *Труды Матем. ин-та АН СССР*, 74, 16—37.

Бояринцев Ю. Е., Узнадзе О. П.

- [1967] Об одном способе решения нестационарного уравнения переноса, *Ж. вычислит. матем. и матем. физ.*, 7 : 6, 1406—1413.

Браиловская И. Ю.

- [1967] Разностный метод численного решения двумерных нестационарных уравнений Навье — Стокса для сжимаемого газа, *Сб. работ ВЦ МГУ*, 7, 3—15.

Браиловская И. Ю., Кускова Т. В., Чудов Л. А.

- [1968] Разностные методы решения уравнений Навье — Стокса, *Сб. работ ВЦ МГУ*, 11, 3—18.

Браиловская И. Ю., Чудов Л. А.

- [1962] Решение уравнений пограничного слоя разностным методом, *Сб. работ ВЦ МГУ*, 1, 167—182.

Будак Б. М.

- [1955] К решению краевых задач параболического типа, *Вестник МГУ*, 8, 33—38.
[1956] О методе прямых для решения некоторых краевых задач, *ДАН СССР*, 109 : 1, 9—12.
[1961] О методе прямых для некоторых квазилинейных краевых задач параболического типа, *Ж. вычислит. матем. и матем. физ.*, 1 : 6, 1105—1112.

- [1962a] О методе прямых для некоторых квазилинейных краевых задач параболического типа с разрывными данными, *Вестник МГУ, сер. матем.*, 3, 3—8.
- [1962b] Об однородных дифференциально-разностных схемах второго порядка точности для параболических и гиперболических уравнений с разрывными коэффициентами, *ДАН СССР*, 142 : 5, 986—989.
- Будак Б. М., Балашова Г. С., Пирмамедов В. Г.
- [1971] О применении разностного метода с выпрямлением фронтов к задаче движения газового объема в водонасыщенной пористой среде, *Труды ВЦ МГУ*, 134—155.
- Будак Б. М., Васильев Ф. П.
- [1963] Сходимость и оценка погрешности метода прямых для решения некоторых двумерных задач фильтрации, *Сб. работ ВЦ МГУ*, 2, 211—238.
- Будак Б. М., Васильев Ф. П., Успенский А. Б.
- [1965] Разностные методы решения некоторых краевых задач типа Стефана, *Сб. работ ВЦ МГУ*, 4, 139—183.
- Будак Б. М., Васильева В. Н.
- [1971] О решении обратной задачи Стефана, *Труды ВЦ МГУ*, 65—86.
- Будак Б. М., Гапоненко Ю. Л.
- [1971] О решении задачи Стефана для квазилинейного параболического уравнения с квазилинейными граничными условиями, *Труды ВЦ МГУ*, 235—312.
- Будак Б. М., Гольдман Н. Л.
- [1971] Разностно-итерационный метод решения оптимальной задачи Стефана для квазилинейного параболического уравнения, *Труды ВЦ МГУ*, 155—179.
- Будак Б. М., Гольдман Н. Л., Егорова А. Т., Успенский А. Б.
- [1967] Метод выпрямления фронтов для решения задач типа Стефана в многомерном случае, *Сб. работ ВЦ МГУ*, 8, 103—120.
- Будак Б. М., Гольдман Н. Л., Успенский А. Б.
- [1966] Разностные схемы с выпрямлением фронтов для решения многофронтовых задач типа Стефана, *ДАН СССР*, 167 : 4, 735—738.
- Будак Б. М., Гольдман Н. Л., Чугаев В. Е.
- [1970] Разностный метод решения внешней краевой задачи для конечного круглого цилиндра с основанием на граничной плоскости, *Сб. работ ВЦ МГУ*, 14, 1—30.
- Будак Б. М., Горбунов А. Д.
- [1957] О разностном методе решения нелинейной задачи Гурса, *ДАН СССР*, 117 : 4, 675—678.
- [1958] Метод прямых для решения одной нелинейной краевой задачи в области с криволинейной границей, *ДАН СССР*, 118 : 5, 858—861.
- [1963] Разностные методы решения задачи Гурса для нелинейных гиперболических уравнений, сб. «Вопросы вычислительной математики и вычислительной техники», М., 67—70.
- Будак Б. М., Искендеров А. Д.
- [1967] Об одном классе обратных краевых задач с неизвестными коэффициентами, *ДАН СССР*, 176 : 1, 20—23.

Будак Б. М., Мадатов Д. А.

- [1971] О решении задачи Стефана с неизвестными коэффициентами, Труды ВЦ МГУ, 180—204.

Будак Б. М., Павлов А. Р.

- [1970] Разностно-итерационный метод решения квазилинейного интегро-дифференциального уравнения параболического типа с квазилинейными граничными условиями на неравномерных сетках, Труды ВЦ МГУ, 1—83.

Будак Б. М., Соловьева Е. Н.

- [1970] Разностно-итерационный метод решения двумерного квазилинейного интегро-дифференциального уравнения с квазилинейными граничными условиями, Сб. работ ВЦ МГУ, 12, 115—123.

Будак Б. М., Соловьева Е. Н., Успенский А. Б.

- [1965] Разностный метод со сглаживанием коэффициента для решения задач Стефана, Ж. вычислит. матем. и матем. физ., 5 : 5, 828—840.

Будак Б. М., Успенский А. Б.

- [1965] Разностный метод решения краевых задач для квазилинейных интегро-дифференциальных уравнений параболического и гиперболического типов с разрывными коэффициентами, Сб. работ ВЦ МГУ, 3, 473—508.

Булеев Н. И.

- [1970] Метод неполной факторизации для решения двумерных и трехмерных разностных уравнений типа диффузии, Ж. вычислит. матем. и матем. физ., 10 : 4, 1042.

Веденеев Е. П., Жидков Н. П.

- [1969] Применение метода регуляризации к дифференцированию функции одного переменного, заданной таблично, Сб. работ ВЦ МГУ, 13, 255—261.

Владимиров В. С.

- [1958] Численное решение кинетического уравнения для сферы, сб. «Вычислит. математика», М., 3—33.

- [1961] Математические задачи односкоростной теории переноса частиц, М., Атомиздат.

Владимиров Л. А.

- [1966] Решение задачи о движении границы раздела двух жидкостей, сб. «Численные методы решения задач матем. физики», М., 267—271.

Гаипова А. Н.

- [1968] Об однородной неявной разностной схеме для решения эволюционного уравнения с фазовыми изменениями, Ж. вычислит. матем. и матем. физ., 8 : 3, 534—543.

Гельфанд И. М.

- [1959] Некоторые задачи теории квазилинейных уравнений, УМН, 14 : 2, 87—158.

Гельфанд И. М., Локуцкий О. В.

- [1962] Метод «прогонки» для решения разностных уравнений, дополнение к книге: Годунов С. К., Рябенский В. С. [1962].

Глаголева Ю. П., Жогов Б. М., Кирьянов Ю. Ф., Мальшаков В. Д., Нестеренко Л. В., Подливаев И. Ф., Софронов И. Д.

- [1971] Основы методики «Медуза» численного расчета двумерных нестационарных задач газодинамики, информ. бюлл. «Численные методы механики сплошной среды», Новосибирск.

Годунов С. К.

- [1962] Разностные методы решения уравнений газовой динамики, Новосибирск.

Годунов С. К., Забродин А. В.

- [1962] О разностных схемах второго порядка точности для многомерных задач, *Ж. вычислит. матем. и матем. физ.*, 2 : 4, 706—708.

Годунов С. К., Семендяев К. А.

- [1962] Разностные методы численного решения задач газовой динамики, *Ж. вычислит. матем. и матем. физ.*, 2 : 1, 3—14.

Годунов С. К., Султангазин У. М.

- [1971] О разностных моделях кинетического уравнения Больцмана, *УМН*, 26 : 3, 3—52.

Гольдин В. Я.

- [1960] Характеристическая разностная схема для нестационарного кинетического уравнения, *ДАН СССР*, 133 : 4, 748—751.

- [1964] Квазидиффузионный метод решения кинетического уравнения, *Ж. вычислит. матем. и матем. физ.*, 4 : 6, 1078—1087.

Гулин А. В.

- [1968] Необходимые и достаточные условия устойчивости трехслойных разностных схем, *Ж. вычислит. матем. и матем. физ.*, 8 : 4, 899—908.

Дмитриев В. И., Захаров Е. В.

- [1969] О численных методах решения некоторого класса дифракционных задач, *Сб. работ ВЦ МГУ*, 13, 139—144.

Дмитриев Н. А., Дмитриева Л. В., Малиновская Е. В., Софронов И. Д., Романцева А. Н.

- [1971] Методика расчета двумерных нестационарных задач гидродинамики в переменных Лагранжа, Информ. бюлл. «Численные методы механики сплошной среды», Новосибирск, изд. СО АН СССР, 15—27.

Дородницын А. А.

- [1958] Об одном численном методе решения некоторых нелинейных задач аэрогидродинамики (метод интегральных соотношений), Труды III Всесоюзного матем. съезда, т. III, 447—453.

- [1960] Об одном методе решения уравнений ламинарного пограничного слоя, *ПМТФ*, 1 : 3, 111—118.

- [1961a] Метод интегральных соотношений, Труды IV Всесоюзного матем. съезда, т. II, 621.

- [1961b] Об одном методе решения уравнений пограничного слоя, сб. «Некоторые проблемы матем. и мех.», Новосибирск, 77—83.

Дьяконов Б. Г.

- [1962a] Разностные схемы с расщепляющимся оператором для нестационарных уравнений, *ДАН СССР*, 144 : 1, 29—32.

- [19626] Метод сеток для решения уравнений параболического типа $2m$ -го порядка с разделяющимися переменными, *ДАН СССР*, 142 : 6, 1236—1238.
- [1963] О применении разностных схем с расщепляющимся оператором для гиперболических уравнений с переменными коэффициентами, *ДАН СССР*, 151 : 4, 762—765.
- [1965] Разностные схемы второго порядка точности с расщепляющимся оператором для многомерных параболических уравнений с переменными коэффициентами, *Сб. работ ВЦ МГУ*, 3, 163—190.
- [1966] О применении разностных схем с расщепляющимся оператором для некоторых систем интегро-дифференциальных уравнений, *Вестник МГУ*, 5, 3—11.
- [1971] О некоторых операторных неравенствах и их применениях, *ДАН СССР*, 198 : 5, 1007—1010.

Дьяченко В. Ф.

- [1961] О численном счете разрывных решений квазилинейных систем, *Ж. вычислит. матем. и матем. физ.*, 1 : 6, 1126—1129.
- [1965] Об одном новом методе численного решения нестационарных задач газовой динамики с двумя пространственными переменными, *Ж. вычисл. матем. и матем. физ.*, 5 : 4, 680—688.

Жидков Н. П., Корнейчук А. А., Крылов А. Л.

- [1963] Численное решение задачи о движении вязкой жидкости между двумя вращающимися цилиндрами, *Сб. работ ВЦ МГУ*, 1, 87—120.

Жуков А. И.

- [1960] Применение метода характеристик к численному решению одномерных задач газовой динамики, *Труды Матем. ин-та АН СССР*, 58, 17—46.

Загускин В. Л., Кондрашов В. Е.

- [1965] О счете уравнений теплопроводности и газовой динамики прогонкой по отдельным областям, *ДАН СССР*, 163 : 5, 1107—1109.

Зуев А. И.

- [1966] О трехслойной схеме для численного интегрирования уравнений газодинамики и нелинейного уравнения теплопроводности, *сб. «Численные методы решения задач математической физики»*, М., 230—236.

Калиткин Н. Н.

- [1968] Прогонка при бесконечном коэффициенте теплопроводности, *Ж. вычисл. матем. и матем. физ.*, 8 : 3, 684—693.

Карманов В. Г.

- [1954] Об одной граничной задаче для уравнения смешанного типа, *ДАН СССР*, 95 : 3, 439—442.
- [1958] О существовании решений некоторых краевых задач для уравнения смешанного типа, *Изв. АН СССР, сер. матем.*, 22 : 1, 117—134.

Коновалов А. Н.

- [1962] Метод дробных шагов решения задачи Коши для многомерного уравнения колебаний, *ДАН СССР*, 147 : 1, 25—27.
- [1964] Применение метода расщепления к численному решению динамических задач теории упругости, *Ж. вычислит. матем. и матем. физ.*, 4 : 4, 760—764.

Коновальцев И. В.

- [1968a] Об устойчивости двухслойных разностных схем, *Ж. вычисл. матем. и матем. физ.*, 8 : 2, 322—333.
- [1968б] Устойчивость в C и L_p двухслойных разностных схем с переменными коэффициентами, *Ж. вычислит. матем. и матем. физ.*, 8 : 4, 894—899.

Котик И. П.

- [1969] Численное исследование волноводного трансформатора, *Сб. работ ВЦ МГУ*, 13, 49—66.

Красносельский М. А., Вайко Г. М., Забрейко П. П., Рутцкий П. П., Стеценко В. Я.

- [1969] Приближенное решение операторных уравнений, М., «Наука».

Кружков С. Н.

- [1961] Об априорной оценке решений линейных параболических уравнений и решении краевых задач для некоторого класса квазилинейных параболических уравнений, *ДАН СССР*, 138 : 5, 1005—1008.
- [1967а] Нелинейные параболические уравнения с двумя независимыми переменными, *Труды Моск. матем. об-ва*, 16, 329—346.
- [1967б] Об оценках старших производных для решений эллиптических и параболических уравнений с непрерывными коэффициентами. *Математ. заметки*, 2 : 5, 549—560.
- [1967в] О некоторых задачах с неизвестной границей для уравнения теплопроводности, *Прикладн. матем. и механ.*, 31 : 5, 1009—1020.
- [1968] Об одном классе задач с неизвестной границей для уравнения теплопроводности, *ДАН СССР*, 178 : 5, 1036—1038.

Крылов В. И., Бобков В. В.

- [1963] О методе интегральных соотношений для задачи Гурса, *ДАН БССР*, 7 : 7, 433—438.
- [1965] Метод интегральных соотношений для уравнений и систем гиперболического типа, *Дифференциальные уравнения*, 1 : 2, 230—248.

Крылов В. И., Лискавец О. А.

- [1963] Оценка погрешности метода прямых для задачи Гурса, *ДАН БССР*, 7 : 8, 505—509.
- [1964] Метод прямых для нестационарных смешанных задач и оценка среднеквадратичной погрешности, *ДАН БССР*, 8 : 6, 353—356.

Кряквина С. А.

- [1966] О точности схем переменных направлений для уравнения теплопроводности, *Ж. вычислит. матем. и матем. физ.*, 6 : 3, 582—589.

Кузнецов Н. Н.

- [1967а] К расчету одномерного течения хорошо проводящего газа в полуограниченной трубе с поршнем, *Сб. работ ВЦ МГУ*, 7, 242—246.
- [1967б] О слабом решении задачи Коши для многомерного квазилинейного уравнения, *Матем. заметки*, 2 : 4, 401—410.
- [1967в] Некоторые математические задачи хромотографии, *Сб. работ ВЦ МГУ*, 6, 242—258.
- [1971] Слабая устойчивость и асимптотическое решение конечноразностных уравнений, *ДАН СССР*, 200 : 5, 18—21.
- [1972] О методе сглаживания, *Сб. работ ВЦ МГУ*, 19, 22—37.

Кузнецов Н. Н., Сторова Е. Н.

- [1967] Одномерное течение проводящего газа в магнитогидродинамической пушке, *Сб. работ ВЦ МГУ*, 7, 247—263.

Куропатенко В. Ф.

- [1966] О разностных методах для уравнений гидродинамики. Труды Матем. ин-та АН СССР, 74, 107—130.

Ладыженская О. А.

- [1951] О применении метода конечных разностей к решению задачи Коши для гиперболических систем, *ДАН СССР*, 88 : 4, 607—610.
[1953] Смешанная задача для гиперболического уравнения, М., ГИТТЛ.
[1957] Метод конечных разностей в теории уравнений с частными производными, *УМН*, 12 : 5, 123—148.

Ладыженская О. А., Головкин К. К.

- [1960] О решениях нестационарной краевой задачи для уравнений Навье — Стокса, Труды Матем. ин-та АН СССР, 59, 100—114.

Ладыженская О. А., К живицки А.

- [1966] Метод сеток для нестационарных уравнений Навье — Стокса. Труды Матем. ин-та АН СССР, 92, 93—99.

Ладыженская О. А., Солонников В. А., Уральцева Н. Н.

- [1967] Линейные и квазилинейные уравнения параболического типа, М., «Наука».

Ладыженская О. А., Уральцева Н. Н.

- [1962] Краевая задача для линейных и квазилинейных уравнений и систем параболического типа, I, II, *Изв. АН СССР, сер. матем.*, 26 : 1, 5—52; 26 : 5, 753—781.
[1963] Краевая задача для линейных и квазилинейных уравнений и систем параболического типа, III, *Изв. АН СССР, сер. матем.*, 27 : 1, 161—240.

Лебедев В. И.

- [1955] Уравнения и сходимость дифференциально-разностного метода (метода прямых), *Вестник МГУ*, 10, 47—57.
[1960] О методе сеток для третьей краевой задачи, *ДАН СССР*, 134 : 2, 267—270.
[1966] Метод характеристик для решения кинетического уравнения, *Ж. вычислит. матем. и матем. физ.*, 6 : 2, 251—275.
[1969] О сходимости КР-метода для некоторых задач переноса, *Ж. вычислит. матем. и матем. физ.*, 9 : 1, 226—235.

Лебедев В. И., Дьяконов Е. Г.

- [1965] О применении разностных схем с расщепляющимся оператором для решения третьей краевой задачи в случае уравнений параболического типа, *Сиб. матем. ж.*, 6 : 1, 108—113.
[1967] Метод расщепления оператора при решении третьей краевой задачи для параболических уравнений, *Сб. работ ВЦ МГУ*, 6, 121—143.

Лискавец О. А.

- [1964] Решение задач для гиперболических уравнений методом прямых. Строгие оценки погрешности, *ДАН БССР*, 8 : 10, 623—626.

- [1965] Метод прямых для одномерных нестационарных смешанных задач и оценка среднеквадратичной погрешности, *Ж. вычислит. матем. и матем. физ.*, 5 : 2, 360—363.

М а л и н о в с к а я Е. В.

- [1969] Применение факторизации к решению кинетического уравнения методом сферических гармоник, сб. «Вычислительные методы в теории переноса», М., Атомиздат.

М а р ч у к Г. И.

- [1967] Численные методы в прогнозе погоды, М., Гидрометеиздат.

М о р о з о в В. А.

- [1962] Сходимость метода прямых для некоторых краевых задач, *Вестник МГУ*, 5, 25—33.
[1964] О дифференциально-разностных схемах второго порядка точности для квазилинейных задач параболического типа с разрывными элементами, *Вестник МГУ*, 2, 12—22.
[1966] Оценка погрешности метода прямых для некоторых краевых задач, *Сб. работ ВЦ МГУ*, 5, 74—86.

М о р о з о в В. Н.

- [1962] О решении кинетических уравнений с помощью КР-метода, сб. «Теория и методы расчета ядерных реакторов», М., Атомиздат.

О л е й н и к О. А.

- [1959] Решение основных краевых задач для уравнений второго порядка с разрывными коэффициентами, *ДАН СССР*, 124 : 6, 1219—1222.
[1960] Об одном методе решения общей задачи Стефана, *ДАН СССР*, 135 : 5, 1054—1057.
[1961] О сходимости некоторых разностных схем, *ДАН СССР*, 137 : 3, 523—526.
[1967а] О решении системы уравнений Прандтля методом конечных разностей, *Прикл. матем. и механ.*, 31 : 1, 90—100.
[1967б] О построении решений системы уравнений пограничного слоя методом прямых, *ДАН СССР*, 174 : 6, 1288—1291.

О л е й н и к О. А., В е н т ц е л ь Т. Д.

- [1957] Первая краевая задача и задача Коши для квазилинейных уравнений параболического типа, *Матем. сборн., нов. сер.*, 41 : 1, 105—128.

О л е й н и к О. А., И л ь и н А. М.

- [1958] О поведении решений задачи Коши для некоторых квазилинейных уравнений при неограниченном возрастании времени, *ДАН СССР*, 120 : 1, 25—28.

О л е й н и к О. А., И л ь и н А. М., К а л а ш н и к о в А. С.

- [1962] Линейные уравнения второго порядка параболического типа, *УМН*, 17 : 3, 3—146.

О л е й н и к О. А., К а л а ш н и к о в А. С., Ч ж о у Ю й - л и н ь

- [1958] Задача Коши и краевые задачи для уравнений типа нестационарной фильтрации, *Изв. Акад. наук СССР*, 22 : 5, 667—764.

О л е й н и к О. А., К р у ж к о в С. Н.

- [1961] Квазилинейные параболические уравнения второго порядка со многими независимыми переменными, *УМН*, 16 : 5, 115—155.

Павлов Б. М.

- [1967] Численное решение задачи Лагранжа при подводе тепла, *Сб. работ ВЦ МГУ*, 7, 264—269.
- [1968] Численное решение задачи о сверхзвуковом вязком течении газа около затупленного тела, *Сб. работ ВЦ МГУ*, 11, 32—43.
- [1969] О расчете сверхзвукового обтекания затупленных тел с использованием полных уравнений Навье — Стокса, *Сб. работ ВЦ МГУ*, 12, 26—38.

Павлов Б. М., Волконская Т. Г.

- [1965] Расчет процесса сжатия в поршневых установках, *Сб. работ ВЦ МГУ*, 4, 184—210.

Павлов Б. М., Петров В. Н.

- [1968] Численное исследование нестационарного теплообмена при движении реального газа, *Сб. работ ВЦ МГУ*, 11, 163—176.

Пасконов В. М.

- [1965] Решение уравнений нестационарного пограничного слоя разностным методом, *Сб. работ ВЦ МГУ*, 4, 130—138.
- [1967] Об одном алгоритме для решения задач пограничного слоя, *Сб. работ ВЦ МГУ*, 7, 150—155.
- [1968] Численное решение нестационарных уравнений пограничного слоя, *Сб. работ ВЦ МГУ*, 11, 75—92.
- [1971] Разностные схемы на самоорганизующемся множестве точек в двумерных односвязных областях произвольной формы, *Ж. вычислит. матем. и матем. физ.*, 11 : 5, 766—782.

Пасконов В. М., Рябинькина Н. В.

- [1965] Решение уравнения нестационарного пограничного слоя разностным методом, *Сб. работ ВЦ МГУ*, 4, 130—138.

Пасконов В. М., Чудов Л. А.

- [1968] Разностные методы расчета течений в пограничном слое, *Сб. работ ВЦ МГУ*, 11, 55—74.

Петровский И. Г.

- [1961] Лекции об уравнениях с частными производными, изд. 3-е, М., Физматгиз.

Положий Г. Н.

- [1962] Численное решение двумерных и трехмерных задач математической физики и функции дискретного аргумента, Киев, «Наукова думка».
- [1966] О некоторых вопросах метода суммарных представлений, *Вычислительная математика*, Киев, 2, 31—52.

Попов Ю. П., Самарский А. А.

- [1969] Полностью консервативные разностные схемы, *Ж. вычислит. матем. и матем. физ.*, 9 : 4, 880—890.
- [1970] Полностью консервативные разностные схемы для уравнений магнитной гидродинамики, *Ж. вычислит. матем. и матем. физ.*, 10 : 4, 990—998.

Рождественский Б. Л., Яненко Н. Н.

- [1968] Системы квазилинейных уравнений, М., «Наука».

Росляков Г. С., Дроздова Н. В.

[1963] Численный расчет обтекания ступенчатого конуса, *Сб. работ ВЦ МГУ*, 1, 18—28.

[1968] Сверхзвуковое обтекание ступенчатых конусов, *Сб. работ ВЦ МГУ*, 11, 123—138.

Росляков Г. С., Костенбойм Х. С.

[1971] Численное решение одномерных задач о взрыве, М., Изд-во МГУ.

Росляков Г. С., Теленин Г. Ф.

[1963] Обзор работ по расчету стационарных осесимметричных течений газа, выполненных в ВЦ МГУ, *Сб. работ ВЦ МГУ*, 2, 5—19.

[1968] Обзор работ по численному исследованию внешних и внутренних задач аэрогидродинамики, выполненных в ВЦ МГУ, *Сб. работ ВЦ МГУ*, 11, 93—112.

Росляков Г. С., Чудов Л. А.

[1971] Некоторые применения метода сеток в газовой динамике, I, II, III, IV, М., Изд-во МГУ.

Русанов В. В.

[1960] Об устойчивости метода матричной прогонки, *сб. «Вычислит. матем.»*, 6, изд. ВЦ АН СССР, 74—83.

[1963] Характеристики общих уравнений газовой динамики, *Ж. вычислит. матем. и матем. физ.*, 3 : 3, 508—527.

Рябенский В. С.

[1967] Спектр семейства разностных операторов над функциями на сеточном графе, *Ж. вычислит. матем. и матем. физ.*, 7 : 6, 1392—1398.

[1970] Расчет теплопроводности на системе стержней, *Ж. вычисл. матем. и матем. физ.*, 10 : 1, 236—247.

[1971] Метод внутренних граничных условий в теории разностных краевых задач, *УМН*, 26 : 3, 105—160.

Самарский А. А.

[1959] О сходимости метода Рунге для уравнения теплопроводности с разрывным коэффициентом теплопроводности. *Научн. докл. высш. школы, физ.-матем. сер.*, 1, 48—53.

[1961a] Априорные оценки для решения разностного аналога дифференциального уравнения параболического типа, *Ж. вычислит. матем. и матем. физ.*, 1 : 3, 441—460.

[1961б] Априорные оценки для разностных уравнений, *Ж. вычислит. матем. и матем. физ.*, 1 : 6, 972—1000.

[1962a] О сходимости и точности однородных разностных схем для одномерных и многомерных параболических уравнений, *Ж. вычислит. матем. и матем. физ.*, 2 : 4, 603—634.

[1962б] Однородные разностные схемы для нелинейных уравнений параболического типа, *Ж. вычислит. матем. и матем. физ.*, 2 : 1, 25—56.

[1962в] О сходимости метода дробных шагов для уравнения теплопроводности, *Ж. вычислит. матем. и матем. физ.*, 2 : 6, 1117—1121.

[1962г] Об одном экономичном разностном методе решения многомерного параболического уравнения в произвольной области, *Ж. вычислит. матем. и матем. физ.*, 2 : 5, 787—811.

[1963a] Локально-одномерные разностные схемы на неравномерных сетках, *Ж. вычислит. матем. и матем. физ.*, 3 : 3, 431—466.

- [1963б] Однородные разностные схемы на неравномерных сетках для уравнений параболического типа, *Ж. вычислит. матем. и матем. физ.*, 3 : 2, 266—298.
- [1963в] Схемы повышенного порядка точности для многомерного уравнения теплопроводности, *Ж. вычислит. матем. и матем. физ.*, 3 : 5, 812—840.
- [1964а] Локально одномерные разностные схемы для многомерных уравнений гиперболического типа, *Ж. вычислит. матем. и матем. физ.*, 4 : 4, 638—648.
- [1964б] Об одной разностной схеме повышенного порядка точности для уравнений теплопроводности с несколькими пространственными переменными, *Ж. вычислит. матем. и матем. физ.*, 4 : 1, 161—165.
- [1964в] Экономичные разностные схемы для систем уравнений параболического типа, *Ж. вычислит. матем. и матем. физ.*, 4 : 5, 927—930.
- [1964г] Экономичные разностные схемы для уравнений параболического типа со смешанными производными, *Ж. вычислит. матем. и матем. физ.*, 4 : 4, 753—759.
- [1965] О монотонных разностных схемах для эллиптических и параболических уравнений в случае несамосопряженного эллиптического оператора, *Ж. вычислит. матем. и матем. физ.*, 5 : 3, 548—551.
- [1967] О регуляризации разностных схем, *Ж. вычислит. матем. и матем. физ.*, 7 : 1, 62—93.
- [1969] Необходимые и достаточные условия для устойчивости двухслойных разностных схем, *ДАН СССР*, 181 : 4, 808—811.
- [1970] Некоторые вопросы общей теории разностных схем, сб. «Дифференциальные уравнения с частными производными», труды Симпозиума, посвященного 60-летию акад. С. Л. Соболева, «Наука», 1023—1036.
- [1971] Введение в теорию разностных схем, М., «Наука».
- Самарский А. А., Гулин А. В.
- [1970] Об устойчивости разностных схем по правым частям, *ДАН СССР*, 192 : 2, 285—288.
- Самарский А. А., Мостинская С. В.
- [1965] Численное решение уравнения теплопроводности с разрывными коэффициентами в полярных координатах, *Сб. работ ВЦ МГУ*, 3, 147—162.
- Самарский А. А., Соболев И. М.
- [1963] Примеры численного расчета температурных волн, *Ж. вычислит. матем. и матем. физ.*, 3 : 4, 702—719.
- Самарский А. А., Фрязинов И. В.
- [1961] О сходимости однородных разностных схем для уравнения теплопроводности с разрывными коэффициентами, *Ж. вычислит. матем. и матем. физ.*, 1 : 5, 806—824.
- [1971] О сходимости локально одномерной схемы решения многомерного уравнения теплопроводности на неравномерных сетках, *Ж. вычислит. матем. и матем. физ.*, 11 : 3, 642—657.
- Самарский А. А., Хао Шоу
- [1967] Однородные разностные схемы на неравномерных сетках для уравнения четвертого порядка, *Сб. работ ВЦ МГУ*, 6, 3—16.
- Саульев В. К.
- [1958] О методах повышенной точности и двусторонних приближениях к решению параболических уравнений, *ДАН СССР*, 118 : 6, 1088—1090.

- [1960] Интегрирование уравнений параболического типа методом сеток, М., Физматгиз.
- [1962а] Об одном методе автоматизации решения краевых задач на быстродействующих вычислительных машинах, *ДАН СССР*, 144 : 3, 497—500.
- [1962б] Об однородных сеточных схемах, *ДАН СССР*, 144 : 4, 723—726.
- [1962в] Пример применения теоремы вложения С. Л. Соболева в вычислительной математике, *ДАН СССР*, 147 : 2, 303—305.
- [1963] О многопараметрических семействах сеточных уравнений для численного интегрирования нестационарных уравнений диффузии, *Ж. вычислит. матем. и матем. физ.*, 3 : 1, 198—201.
- [1965] О регулярных неоднородных сеточных схемах и квазистойчивости, *Дифференциальные уравнения*, 1 : 3, 421—425.

С в е ш н и к о в А. Г.

- [1962] К расчету согласования плоских волноводов, *Ж. вычислит. матем. и матем. физ.*, 2 : 1, 175—179.
- [1963а] К обоснованию метода расчета нерегулярных волноводов, *Ж. вычислит. матем. и матем. физ.*, 3 : 1, 170—179.
- [1963б] Обоснование методов исследования распространения электромагнитных колебаний в волноводах с анизотропным заполнением, *Ж. вычислит. матем. и матем. физ.*, 3 : 5, 953—955.
- [1966] Обоснование методов исследования распространения колебаний в нерегулярных волноводах, Тезисы крат. научн. сообщений Международного конгресса математиков, секция 12, 49.
- [1969] Дифракция на звездном теле в среде с переменными коэффициентами (плоский случай), *Сб. работ ВЦ МГУ*, 13, 145—151.

С в е ш н и к о в А. Г., В о л к о в Б. И., Е ф и м о в В. В., С е м а ш к о Н. Н.

- [1971] Расчет движения пучка заряженных частиц в электростатическом поле с учетом пространственного заряда, *Сб. работ ВЦ МГУ*, 16, 244—263.

С в е ш н и к о в А. Г., И л ь и н с к и й А. С.

- [1969] Прямые методы исследования волноводных систем, *Сб. работ ВЦ МГУ*, 13, 3—26.
- [1969] Метод исследования плоских волноводов с импедансными граничными условиями и резким изменением боковой поверхности, *Сб. работ ВЦ МГУ*, 13, 27—33.

С в е ш н и к о в А. Г., И л ь и н с к и й А. С., К р а в ц о в В. В.

- [1969] Рассеяние электромагнитных волн в полых направляющих системах, *Сб. работ ВЦ МГУ*, 13, 56—71.

С в е ш н и к о в А. Г., И л ь и н с к и й А. С., П а в л о в А. Л.

- [1971] Численное решение задачи дифракции на неоднородном ограниченном теле, *Сб. работ ВЦ МГУ*, 16, 116—124.

С в е ш н и к о в А. Г., С а м а р с к и й А. А., З а х а р о в В. И.

- [1971] Применение метода больших частиц к расчету движения заряженного пучка в электромагнитном поле с учетом пространственного заряда пучка, *Сб. работ ВЦ МГУ*, 16, 225—243.

С в е ш н и к о в А. Г., С е м а ш к о Н. Н., Х а п а е в М. М.

- [1971] О преобразовании длинных ленточных пучков заряженных частиц в цилиндрические, *Сб. работ ВЦ МГУ*, 16, 264—271.

Сидоров А. Ф.

- [1966] Об одном алгоритме расчета оптимальных разностных схем, Труды Матем. ин-та АН СССР, 74, 147—151.

Софонов И. Д.

- [1965] Разностная схема с диагональными направлениями прогонок для решения уравнения теплопроводности, Ж. вычислит. матем. и матем. физ., 5 : 2, 347—350.

Тихонов А. Н., Бабенко К. И.

- [1963] О разностных методах решения задач математической физики, Труды 4-го Всесоюзного матем. съезда, 1, 243.

Тихонов А. Н., Дмитриев В. И.

- [1969] О методах решения обратной задачи теории антенн, Сб. работ ВЦ МГУ, 13, 209—214.

Тихонов А. Н., Самарский А. А.

- [1958] Об однородных разностных схемах, ДАН СССР, 122 : 4, 562—565.
[1959а] О сходимости разностных схем в классе разрывных коэффициентов, ДАН СССР, 124 : 3, 529—532.
[1959б] Об одной наилучшей однородной разностной схеме, ДАН СССР, 124 : 4, 779—782.
[1960а] О канонических однородных разностных схемах, ДАН СССР, 131 : 4, 761—764.
[1960б] Коэффициентно устойчивые разностные схемы, ДАН СССР, 131 : 6, 1264—1267.
[1960в] Об однородных разностных схемах высокого порядка точности, ДАН СССР, 131 : 3, 514—517.
[1961а] Об однородных разностных схемах, Ж. вычислит. матем. и матем. физ., 1 : 1, 5—63.
[1961б] Разностная задача Штурма — Лиувилля, Ж. вычислит. матем. и матем. физ., 1 : 5, 785—805.
[1961в] Однородные разностные схемы высокого порядка точности на неравномерных сетках, Ж. вычислит. матем. и матем. физ., 1 : 3, 425—440.
[1962] Однородные разностные схемы на неравномерных сетках, Ж. вычислит. матем. и матем. физ., 2 : 5, 812—832.
[1963а] К теории однородных разностных схем. Материалы к совместному советско-американскому Симпозиуму по уравнениям с частными производными, Новосибирск, 3—10.
[1963б] Об однородных разностных схемах высокого порядка точности на неравномерных сетках, Ж. вычислит. матем. и матем. физ., 3 : 1, 99—108.
[1963в] Об устойчивости разностных схем, ДАН СССР, 149 : 3, 529—531.
[1966] Уравнения математической физики, изд. 3-е, М., «Наука».

Тихонов А. Н., Самарский А. А. и др.

- [1969] Расчет ионизационной неустойчивости в низкотемпературной намагниченной плазме, ДАН СССР, 184 : 3, 452—453.

Трощнев В. Е.

- [1966] Решение кинетического уравнения и уравнений квазидиффузии по согласованным разностным схемам, сб. «Численные методы решения задач математической физики», М., «Наука», 177—185.

- [1970] О дивергентности схемы «крест» численного решения уравнений газовой динамики, информ. бюлл. «Численные методы механики сплошной среды», Новосибирск, изд. СО АН СССР (сер. техн. наук), 1 : 5, 87—96.

Филиппов А. Ф.

- [1955] Об устойчивости разностных уравнений, *ДАН СССР*, 100 : 5, 1045—1048.

Филиппов А. Ф., Рябенский В. С.

- [1956] Об устойчивости разностных уравнений, М., Физматгиз.

Франк Л. С.

- [1964] Разностные методы решения задачи Коши для некоторых систем первого порядка, *УМН*, 19 : 5, 189—192.
[1968] Разностные методы решения некорректной задачи Коши, моделирующей течение идеального газа в сопле, *ДАН СССР*, 182 : 3, 526—529.
[1969] О гиперболических разностных операторах, *ДАН СССР*, 189 : 3, 486—488.

Франкль Ф. И., Алексеева Р. Н.

- [1934] Две краевые задачи из теории гиперболических уравнений в частных производных с приложением к сверхзвуковым газовым течениям, *Матем. сб.*, 41 : 4, 483—502.

Фрязинов И. В.

- [1964] О разностной аппроксимации граничных условий для третьей краевой задачи, *Ж. вычисл. матем. и матем. физ.*, 4 : 6, 1106—1112.
[1968] Экономичные симметризованные разностные схемы решения краевых задач для многомерных уравнений параболического типа, *Ж. вычислит. матем. и матем. физ.*, 8 : 2, 436—443.
[1969а] Априорные оценки для одного семейства экономичных схем, *Ж. вычислит. матем. и матем. физ.*, 9 : 3, 595—604.
[1969б] Экономичные разностные схемы повышенного порядка точности для решения многомерного уравнения параболического типа, *Ж. вычислит. матем. и матем. физ.*, 9 : 6, 1316—1326.

Христианович С. А.

- [1937] Задача Cauchy для нелинейных уравнений гиперболического типа, *Матем. сб.*, 2 : 7, 871—900.

Христианович С. А., Левин Л. И., Павлов Е. П.

- [1943] О расчете сопел Лавалья, *Прикл. матем. и механика*, 7 : 1, 3—24.

Чудов Л. А.

- [1967] Разностные методы и некорректные задачи для уравнений с частными производными, *Сб. работ ВЦ МГУ*, 8, 34—62.

Чудов Л. А., Воробьева Н. А., Франк Л. С. и др.

- [1968] Разностные методы решения некорректной задачи Коши для трехмерного уравнения Лапласа, *Сб. работ ВЦ МГУ*, 11, 147—155.

Чудов Л. А., Костенбойм Х. С., Турецкая А. Т.

- [1969] Точечный взрыв в неоднородной атмосфере, *Ж. прикл. мех. и тех. физ.*, 5, 25—28.

Чудов Л. А., Морозова Т. К.

- [1965] Разностный метод решения о движении газа в трубе переменного сечения, *Сб. работ ВЦ МГУ*, 4, 232—241.

Чушкин П. И.

- [1968] Метод интегральных соотношений для сверхзвуковых пространственных течений, *Ж. вычислит. матем. и матем. физ.*, 8 : 4, 853—864.

Чушкин П. И., Кацкова О. Н.

- [1964] Об одной схеме численного метода характеристик, *ДАН СССР*, 154 : 1, 26—29.

Чушкин П. И., Коробейников П. И.

- [1964] Метод численного расчета взрыва в газах, *ДАН СССР*, 154 : 3, 549—552.
[1968] Плоский, цилиндрический и сферический взрыв в газе с противодействием (численный метод), *Труды Матем. ин-та АН СССР*, 87, 4—34.

Щедрин Б. М., Жидков Н. П.

- [1969] Применение метода материальной точки для определения атомных структур некоторых кристаллов, *Сб. работ ВЦ МГУ*, 13, 262—271.

Яненко Н. Н.

- [1959] Об одном разностном методе счета многомерного уравнения теплопроводности, *ДАН СССР*, 125 : 6, 1207—1210.
[1960] Об экономичных неявных схемах (метод дробных шагов), *ДАН СССР*, 134 : 5, 1034—1036.
[1961] О неявных разностных методах счета многомерного уравнения теплопроводности, *Изв. высш. учебн. завед., сер. матем.*, 4, 148—157.
[1962] О сходимости метода расщепления оператора для уравнения теплопроводности с переменными коэффициентами, *Ж. вычисл. матем. и матем. физ.*, 2 : 5, 933—937.
[1964] О слабой аппроксимации систем дифференциальных уравнений, *Сиб. матем. журн.*, 5 : 6, 1431—1434.
[1967] Метод дробных шагов решения многомерных задач математической физики, Новосибирск.

Яненко Н. Н., Бояринцев Ю. Е.

- [1961] О сходимости разностных схем для уравнения теплопроводности с переменными коэффициентами, *ДАН СССР*, 139 : 6, 1322—1324.

Яненко Н. Н., Сучков В. А., Погодин Ю. Я.

- [1959] О разностном решении уравнения теплопроводности в криволинейных координатах, *ДАН СССР*, 128 : 5, 903—905.

Яненко Н. Н., Фролов В. Д., Неуважаев В. Е.

- [1967] О применении метода расщепления для численного расчета движения теплопроводного газа в криволинейных координатах, *Изв. СО АН СССР, сер. техн. наук*, 8 : 2, 1—30.

Яненко Н. Н., Шокин Ю. И.

- [1968a] О связи корректности первых дифференциальных приближений и устойчивости разностных схем для гиперболических систем уравнений, *Матем. заметки*, 4 : 5, 493—502.
- [1968b] Об аппроксимационной вязкости разностных схем, *ДАН СССР*, 182 : 2, 280—281.
- [1969] О первых дифференциальных приближениях разностных схем для гиперболических систем уравнений, *Сиб. матем. журн.*, 10 : 5, 1173—1187.

Яненко Н. Н., Яушев И. К.

- [1966] Об одной абсолютно устойчивой схеме интегрирования уравнений гидродинамики, *Труды Матем. ин-та АН СССР*, 74, 141—146.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

аппроксимация устойчивая 54

банахово пространство 40

векторное пространство 40
внешняя точка сетки 142
внутренняя точка сетки 142
волновое уравнение 260
вперед 26

главный член ошибки 134
гладкое разложение единицы 115
граничная точка сетки 142

двухслойная формула 52
двухшаговая схема Лакса — Вендроффа 299
двухшаговый метод Лакса — Вендроффа 355
дивергентная форма уравнений гидродинамики 299
дискретно-ординатный метод 236
диссипативная разностная схема 118
достаточное условие Като 90

замкнутое множество 41
замкнутый оператор 47
запаздывающая характеристика 308

инвариант Римана 308
инвариантное подпространство 97
индекс собственного значения 97
инфинитезимальный оператор 58

корректно поставленная задача 49, 105
— — относительно начального элемента задача 351
критерий устойчивости Баченана 90
— — Годунова — Рябенского 160, 166—167
кусочная аналитичность 363
кучная последовательность 89

линейное преобразование 43
— — ограниченное 43
— пространство 40

матрица перехода 76
метод Вика — Чандрасекхара 235
— — — обобщенный 243
— Годунова 334
— дробных шагов 218
— прямого интегрирования 247
— расщепления 218
— Саульева 195
— сферических гармоник 225
— фон Неймана — Рихтмайера 329?
— чередующихся направлений 213
минимально полуограниченный оператор 154
многослойная формула 52
— — устойчивая 178
многошаговая формула 52
множитель перехода 76
модифицированное условие фон Неймана 269

назад 26
неустойчивость 18
неявная система 27
норма линейного преобразования 43
нормированное пространство 40
носитель 73

область значений 42
область определения 42
обобщенное решение 62, 64
обобщенный метод Вика — Чандрасекхара 243
— разрешающий оператор 50
ограниченное линейное преобразование 43
одношаговая формула 52
оператор 42
— замкнутый 47
— инфинитезимальный 58
— минимально полуограниченный 154
— полуограниченный 154
— разрешающий обобщенный 50
операция 42
опережающая характеристика 307—308

первая вариация оператора 135
плотное множество 41
погрешность аппроксимации 30, 53
— метода 19
подгонка скачка 306
полное пространство 40

полугруппа 50
 полуограниченный оператор 154
 порядок точности 76, 181
 предельная точка 41
 преобразование 42
 — линейное 43
 — — ограниченное 43
 принцип равномерной ограниченности 44

равномерная ограниченность 44
 разрешающий оператор обобщенный 50
 расширение 43
 регулярная точка спектра семейства операторов 164
 резольвентное условие 83
 решение обобщенное 62, 64
 — точное 49, 62

сильно устойчивая разностная схема 108
 скалярное произведение 70
 согласованно аппроксимирует 53
 спектр семейства операторов 159
 схема Карлсона 241
 — Кранка — Николсона 196
 сходящаяся аппроксимация 53, 178

теорема Канторовича 46
 — Крайса 105, 118
 — — о матрицах 82
 — Крайса — Стрэнга 67
 — Лакса — Ниренберга 129
 — Лакса об эквивалентности 54
 — Стрэнга 172
 — Фридрихса 128
 — Шура 86
 точка спектра семейства операторов 159
 — типа А 375

точка типа В 376
 точное решение 49, 62

уравнение переноса 222
 уравнения в характеристической форме 369
 — Гюгоньо 306
 — Лагранжа 291
 — — разностные 294
 — Лакса — Вендроффа 301
 — Ренкина — Гюгоньо 306
 — Эйлера 288
 — — разностные 289
 условие согласованности 53, 178
 — Лакса — Вендроффа 98
 — фон Неймана 79
 устойчивая аппроксимация 54
 — многослойная разностная формула 178
 устойчивость 19

фронт 377
 фундаментальная последовательность 40

характеристика 307
 — запаздывающая 308
 — опережающая 307—308
 характеристическая плоскость 369
 хорошо обусловленная система 164

эквивалентные нормы 66
 энергетическое неравенство 154

явная система 27

LW-метод 355
 $p \times p$ -устойчивость 82
 S_n -метод Карлсона 244

ОГЛАВЛЕНИЕ

От редакторов перевода	5
Предисловие ко второму изданию	7
Предисловие к первому изданию	9
ЧАСТЬ I. ОБЩАЯ ТЕОРИЯ	13
Глава 1. Введение	13
§ 1.1. Краевые задачи	13
§ 1.2. Уравнение теплопроводности	14
§ 1.3. Конечноразностные уравнения	16
§ 1.4. Устойчивость	19
§ 1.5. Неявные разностные уравнения	26
§ 1.6. Погрешность аппроксимации	29
§ 1.7. Скорость сходимости	31
§ 1.8. Замечания о формулах высших порядков и погрешностях округ- ления	33
§ 1.9. Содержание следующих глав	34
Глава 2. Линейные операторы	37
§ 2.1. Краевые задачи и функциональные пространства	37
§ 2.2. Банаховы пространства	39
§ 2.3. Линейные операторы в банаховом пространстве	42
§ 2.4. Теорема о расширении оператора	43
§ 2.5. Принцип равномерной ограниченности	44
§ 2.6. Основная теорема о сходимости	46
§ 2.7. Замкнутые операторы	47
Глава 3. Линейные разностные уравнения	48
§ 3.1. Корректно поставленные краевые задачи	48
§ 3.2. Конечноразностные аппроксимации	51
§ 3.3. Сходимость	53
§ 3.4. Устойчивость	54
§ 3.5. Теорема Лакса об эквивалентности	54
§ 3.6. Замкнутый оператор A'	58
§ 3.7. Неоднородные задачи	61
§ 3.8. Изменение нормы	66
§ 3.9. Устойчивость и возмущения	67
Глава 4. Линейные краевые задачи с постоянными коэффициентами	69
§ 4.1. Класс задач	69
§ 4.2. Ряды и интегралы Фурье	70
§ 4.3. Корректно поставленные краевые задачи	72
§ 4.4. Конечноразностные уравнения	74
§ 4.5. Порядок точности и условие согласованности	76
§ 4.6. Устойчивость	77
§ 4.7. Условие фон Неймана	79
§ 4.8. Одно простое достаточное условие	80
§ 4.9. Теорема Крайса о матрицах	81
§ 4.10. Критерий устойчивости Ваченана	89
§ 4.11. Дальнейшие примеры достаточных условий	92
Глава 5. Линейные задачи с переменными коэффициентами. Нелинейные задачи	100
§ 5.1. Введение	100
§ 5.2. Другие определения устойчивости	104

§ 5.3. Параболические уравнения	109
§ 5.4. Диссипативные разностные схемы для симметричных гиперболических уравнений	117
§ 5.5. Дальнейшие результаты для симметричных гиперболических уравнений	128
§ 5.6. Нелинейные уравнения с гладкими решениями	133
Глава 6. Смешанные краевые задачи	139
§ 6.1. Введение	139
§ 6.2. Основные идеи энергетического метода	140
§ 6.3. Простейшие примеры применения энергетического метода: выбор устойчивых аппроксимаций граничных условий и нелинейных членов	145
§ 6.4. Одновременное распространение звука и тепла	150
§ 6.5. Смешанные задачи для симметричных гиперболических систем	153
§ 6.6. Спектральный анализ и критерий устойчивости Годунова — Рябенского	158
§ 6.7. Применение критерия Годунова — Рябенского к смешанным задачам	162
§ 6.8. Заключение	171
Глава 7. Многослойные разностные уравнения	174
§ 7.1. Обозначения	174
§ 7.2. Вспомогательное банахово пространство	175
§ 7.3. Теорема об эквивалентности	178
§ 7.4. Согласованность и порядок точности	181
§ 7.5. Пример Дюфора и Франкела	183
§ 7.6. Резюме	185
ЧАСТЬ II. ПРИЛОЖЕНИЯ	186
Предисловие к части II	186
Глава 8. Диффузия и теплопроводность	188
§ 8.1. Примеры диффузии	188
§ 8.2. Простейшая задача теплопроводности	189
§ 8.3. Переменные коэффициенты	196
§ 8.4. Влияние на устойчивость членов низшего порядка	198
§ 8.5. Решение неявных уравнений	201
§ 8.6. Нелинейные задачи	204
§ 8.7. Задачи с несколькими пространственными переменными	208
§ 8.8. Методы чередующихся направлений	213
§ 8.9. Методы расщепления и дробных шагов	218
Глава 9. Уравнение переноса	220
§ 9.1. Физические основы	220
§ 9.2. Общее уравнение переноса нейтронов	221
§ 9.3. Однородный слой; одна группа	224
§ 9.4. Однородная сфера; одна группа	225
§ 9.5. Метод сферических гармоник	225
§ 9.6. Слой; система разностных уравнений I для гиперболических уравнений	230
§ 9.7. Парадокс	232
§ 9.8. Слой; система разностных уравнений II (схема Фридрихса)	233
§ 9.9. Неявные схемы	234
§ 9.10. Метод Вика — Чандрасекхара для слоя	235
§ 9.11. Эквивалентность двух методов	236

§ 9.12. Граничные условия	238
§ 9.13. Разностные системы I и II	239
§ 9.14. Система III; пространственные разности вперед и назад	239
§ 9.15. Система IV (неявная)	240
§ 9.16. Система V (схема Карлсона)	241
§ 9.17. Обобщение метода Вика — Чандрасекхара	243
§ 9.18. S_n -метод Карлсона [1953]	244
§ 9.19. Метод прямого интегрирования	247
Глава 10. Звуковые волны	259
§ 10.1. Физические основы	259
§ 10.2. Обычное конечноразностное уравнение	260
§ 10.3. Неявная система	262
§ 10.4. Одновременное распространение звука и тепла	263
§ 10.5. Практический критерий устойчивости	268
Глава 11. Упругие колебания	270
§ 11.1. Поперечные колебания тонкого стержня	270
§ 11.2. Явные разностные уравнения	272
§ 11.3. Неявная система	273
§ 11.4. Достоинства неявной системы	274
§ 11.5. Решение неявных уравнений произвольного порядка	274
§ 11.6. Колебания продольно напряженного стержня	281
Глава 12. Одномерное движение жидкости (газа)	287
§ 12.1. Введение	287
§ 12.2. Уравнения Эйлера	288
§ 12.3. Разностные уравнения Эйлера	289
§ 12.4. Уравнения Лагранжа	291
§ 12.5. Разностные уравнения Лагранжа	294
§ 12.6. Поверхности раздела в лагранжевых координатах	297
§ 12.7. Дивергентная форма уравнений гидродинамики и уравнения Лакса — Вендроффа	299
§ 12.8. Условия на скачке	304
§ 12.9. «Подгонка» скачка	306
§ 12.10. Влияние диссипации	309
§ 12.11. Конечноразностные уравнения	315
§ 12.12. Устойчивость конечноразностных уравнений	318
§ 12.13. Численная проверка метода псевдовязкости	321
§ 12.14. Метод Лакса — Вендроффа для расчета движения скачков	327
§ 12.15. Метод Годунова	334
§ 12.16. Магнитная гидродинамика	340
Глава 13. Многомерное движение жидкости (газа)	346
§ 13.1. Введение	346
§ 13.2. Уравнения многомерной гидродинамики	349
§ 13.3. Корректно и некорректно поставленные задачи	351
§ 13.4. Двухшаговый метод Лакса — Вендроффа, или LW-метод	355
§ 13.5. Вязкость в LW-методе	359
§ 13.6. Кусочно-аналитические краевые задачи	362
§ 13.7. Программа развития методов	367
§ 13.8. Характеристики в двумерных течениях	368
§ 13.9. Подгонка скачков в двумерных задачах	371
§ 13.10. Задача о движении атмосферного фронта	377
Библиография	381
Дополнительная библиография	396
Предметный указатель	414

Р. Рихтмайер, К. Мортон

**РАЗНОСТНЫЕ МЕТОДЫ РЕШЕНИЯ
КРАЕВЫХ ЗАДАЧ**

Редакторы В. И. Авербух и Г. М. Ильичева

Художник Л. Г. Ларский

Художественный редактор В. И. Шаповалов

Технический редактор Е. Н. Лебедева

Корректор О. К. Румянцева

Сдано в набор 20/X 1971 г. Подписано к печати 28/IV 1972 г. Бумага
кн.-журн. $60 \times 90^{1/16} = 13,13$ бум. л. 26,25 печ. л. Уч.-изд. л. 24,86
Изд. № 1/6435. Цена 1 р. 93 к. Зак. 1300

ИЗДАТЕЛЬСТВО «МИР»

Москва, 1-й Рижский пер., 2

Ордена Трудового Красного Знамени

Ленинградская типография № 2

имени Евгении Соколовой

Главполиграфпрома Комитета по печати

при Совете Министров СССР

Измайловский проспект, 29